



Optimal control of state constrained age-structured problems

Joseph Frédéric Bonnans, Justina Gianatti

► To cite this version:

Joseph Frédéric Bonnans, Justina Gianatti. Optimal control of state constrained age-structured problems. SIAM Journal on Control and Optimization, 2020, 58 (4), pp.2206–2235. hal-02164310v2

HAL Id: hal-02164310

<https://inria.hal.science/hal-02164310v2>

Submitted on 19 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

OPTIMAL CONTROL OF STATE CONSTRAINED AGE-STRUCTURED PROBLEMS

J. FRÉDÉRIC BONNANS AND JUSTINA GIANATTI

ABSTRACT. The aim of this work is to study an optimal control problem with state constraints where the state is given by an age-structured, abstract parabolic differential equation. We prove the existence and uniqueness of solution for the state equation and provide first and second parabolic estimates. We analyze the differentiability of the cost function and, based on the general theory of Lagrange multipliers, we give a first order optimality condition. We also define and analyze the regularity of the costate.

Keywords. Age-structured systems, optimal control, abstract parabolic equations, state constraints

AMS subject classifications. 49K20, 49K27, 48N60, 92D25, 35K90

1. INTRODUCTION

In the last decades, there were many contributions on the study of age-structured systems. They appear naturally in several applications, in particular to model population dynamics, like it was introduced in [14, 13]. The existence, uniqueness, regularity and stability of solution for different age-structured systems with and without diffusion have been extensively studied, see for instance [17, 16, 12, 7, 22] and the references therein.

There are also several works on optimal control of age-structured systems. Optimality conditions of Pontryagin maximum type are given in [6] and [11], where they consider dynamics without diffusion. For a particular biological application, optimality conditions and numerical results are given in [9]. In the case of dynamics with diffusion, optimality conditions as well as existence of optimal controls are studied in [1, 2, 20], where the control has different meanings depending on the model. The bibliography is vast and we only mention some works, where interested references can be founded.

However, in the case of optimal control problems with age-structured dynamics and state constraints, there are only a few works. In [10] the author deals with a Hamilton Jacobi equation with state constraints and an economical application with an age-structured equation is presented. In [15] the authors study the Pontryagin Maximum Principle for a particular kind of problems and as an illustrative example, they present an optimal control problem for an age-structured system with pointwise terminal state constraints. Finally, in [19] the authors consider mixed state-control constraints, and derive optimality conditions based on the Lagrange multipliers theory.

In this work we address an optimal control problem of partial differential equations, where the state equation is age-structured, and is driven by a general abstract parabolic operator. Most of the population dynamic models considered in the above mentioned works, are particular cases of our formulation. We also deal with a finite number of linear state constraints, and to the best of our knowledge, this was not considered in the literature but seems to arise naturally in some applications.

Throughout this work we consider different types of initial conditions with respect to the age variable. We start by considering a given initial condition, as might occur in a fishery or harvesting problem, and then we study an initial condition given by a fecundity rate, which is a natural birth condition for many population dynamics. Furthermore, we assume

Date: April 18, 2020.

The first author acknowledges support from the Labex Mathématiques Hadamard and iCODE (IDEX Paris-Saclay) for the PGMO project “Optimal control of conservation equations”.

that the cost function includes final costs with respect the final time and final age. This kind of cost again makes sense for the problems mentioned above.

Finally, we tackle the problem of existence of optimal controls. Unfortunately, for this general framework is not easy to obtain such results. In [1] the authors prove the existence of optimal solutions but under a restrictive assumption about the state and the control spaces. For this reason, in the last part of this work we introduce a modified optimal control problem, which is related to the original one, and for which we can prove the existence of optimal controls.

This paper is organized as follows: in Section 2 we present the optimal control problem. We start studying the existence and uniqueness of solution for the state equation in Section 3 and we provide first and second parabolic estimates for this equation. In Section 4 we analyze the differentiability of the cost function and the state constraints, and provide optimality conditions based on the Lagrange multipliers theory. We define the costate and study its regularity. In Section 5 we study a modified optimal control problem where the existence of optimal solutions can be proven.

2. OPTIMAL CONTROL PROBLEM

Let (V, H, V^*) be a *separable Gelfand triple*. We denote by $(\cdot, \cdot)_H$ the scalar product in H and by $\langle \cdot, \cdot \rangle_V$ the duality product between V^* and V .

Let $0 < A < T < \infty$ and $D := (0, T) \times (0, A)$. Given a Banach space U , we define the set of controls as

$$(1) \quad \mathcal{U} := L^\infty(D; U).$$

In this section we introduce the optimal control problem considered in this work:

$$(2) \quad \min J(u, y) := \int_0^T \int_0^A \ell(u, y) da dt + \int_0^A h_0(y(T, a)) da + \int_0^T h_1(y(t, A)) dt$$

s.t.

$$(3) \quad \begin{cases} (y_t + y_a)(t, a) + \mathcal{A}(t, a, u(t, a))y(t, a) = f(t, a), & \text{in } L^2(D; V^*), \\ y(0, a) - y_0(a) = 0, & \text{in } L^2(0, A; H), \\ y(t, 0) - y_1(t) = 0, & \text{in } L^2(0, T; H), \\ \int_0^A (\xi_i(a), y(t, a))_H da - M_i \leq 0, & t \in [0, T], \quad i = 1, \dots, r, \end{cases}$$

where $y(t, a) \in V$ denotes the state function and $u \in \mathcal{U}$ the control. We assume $\ell : H \times H \rightarrow \mathbb{R}$, $h_0, h_1 : H \rightarrow \mathbb{R}$, $f \in L^2(D; V^*)$, $\mathcal{A} : D \times U \rightarrow L(V, V^*)$, $y_0 \in L^2(0, A; H)$, $y_1 \in L^2(0, T; H)$ and $\xi_i \in L^2(0, A; H)$, $i = 1, \dots, r$. Additional assumptions over the data will be provided in the next sections.

Throughout this work we will also consider age-structured state systems where the birth process is depending on a fecundity rate, i.e.

$$(4) \quad y(t, 0) = \int_0^A c(t, a) y(t, a) da,$$

where $c : D \rightarrow \mathbb{R}^+$ is a given function.

3. AGE-STRUCTURED STATE SYSTEM

In this section we study the uncontrolled state equation. We introduce the state space and prove the existence and uniqueness of solution. Under additional assumptions we provide a second parabolic estimate for this case. Finally, we show how our approach can cover the case when the birth process $y(\cdot, 0)$ depends on a fecundity rate of the population.

We begin by studying the following abstract parabolic age-structured system.

$$(5) \quad \begin{cases} (y_t + y_a)(t, a) + \mathcal{A}(t, a)y(t, a) = f(t, a) & \text{in } L^2(D; V^*), \\ y(0, a) = y_0(a) & \text{in } L^2(0, A; H), \\ y(t, 0) = y_1(t) & \text{in } L^2(0, T; H), \end{cases}$$

where $\mathcal{A} : D \rightarrow L(V, V^*)$ is a measurable map, $f \in L^2(D; V^*)$, $y_0 \in L^2(0, A; H)$ and $y_1 \in L^2(0, T; H)$.

We make the following assumptions for \mathcal{A} of uniform continuity and semicoercivity:

- (H1) There exists $C > 0$ such that $\|\mathcal{A}(t, a)v\|_{V^*} \leq C\|v\|_V$, for all $v \in V$ and a.a. $(t, a) \in D$.
- (H2) There exist $\alpha > 0$ and $\lambda \geq 0$ such that $\langle \mathcal{A}(t, a)v, v \rangle_V \geq \alpha\|v\|_V^2 - \lambda\|v\|_H^2$, for all $v \in V$ and a.a. $(t, a) \in D$.

3.1. Space $W(D)$. In this section we introduce the state space.

We recall that if $y \in H$ and $v \in V$ then $\langle y, v \rangle_V = (y, v)_H$, when y is considered as an element of V^* .

We denote by $C_0^\infty(D; V)$ the space of infinitely differentiable functions with compact supports in D and image in V . In particular when the image space is \mathbb{R} we write $C_0^\infty(D)$. Given $y \in L^2(D; V)$, if there exists $w \in L^2(D; V^*)$ such that

$$(6) \quad - \int_0^T \int_0^A (y, \varphi_t + \varphi_a)_H \, da \, dt = \int_0^T \int_0^A \langle w, \varphi \rangle_V \, da \, dt, \quad \forall \varphi \in C_0^\infty(D; V),$$

we say that $w = y_t + y_a$ in a weak sense (which does not imply that y_t and y_a exist in $L^2(D; V^*)$). Since the space V is separable, there exists a countable Hilbert basis, and we can conclude that (6) is equivalent to

$$(7) \quad - \int_0^T \int_0^A (\varphi_t + \varphi_a)(y, v)_H \, da \, dt = \int_0^T \int_0^A \varphi \langle w, v \rangle_V \, da \, dt, \quad \forall (\varphi, v) \in C_0^\infty(D) \times V.$$

The last condition is also equivalent to the following equality in V^* ,

$$(8) \quad - \int_0^T \int_0^A (\varphi_t + \varphi_a)y \, da \, dt = \int_0^T \int_0^A \varphi w \, da \, dt, \quad \forall \varphi \in C_0^\infty(D).$$

Definition 3.1. Given a separable Gelfand triple (V, H, V^*) , we define the space

$$(9) \quad W(D) := \{y \in L^2(D; V) : y_t + y_a \in L^2(D; V^*)\},$$

endowed with the norm

$$(10) \quad \|y\|_{W(D)} := \left(\int_0^T \int_0^A \|y(t, a)\|_V^2 + \|(y_t + y_a)(t, a)\|_{V^*}^2 \, da \, dt \right)^{\frac{1}{2}}.$$

The next result is easy to obtain by following the lines of [18].

Proposition 3.2. *The space $C^\infty(D; V)$ is a dense subset of $W(D)$.*

By the previous proposition and [18], we obtain the following result that appears in [12, Lemma 0] and [16, Section 4].

Lemma 3.3. *Let $y \in W(D)$, then for all $\bar{t} \in [0, T]$ and $\bar{a} \in [0, A]$ there exist the traces*

$$(11) \quad y(\bar{t}, \cdot) \in L^2(0, A; H) \quad \text{and} \quad y(\cdot, \bar{a}) \in L^2(0, T; H),$$

and the maps $\bar{t} \mapsto y(\bar{t}, \cdot)$ and $\bar{a} \mapsto y(\cdot, \bar{a})$ are continuous. Also, the maps $W(D) \rightarrow C(0, T; L^2(0, A; H))$ and $W(D) \rightarrow C(0, A; L^2(0, T; H))$ are continuous.

In addition, the following integration by parts formula holds, for all $0 \leq t_1 \leq t_2 \leq T$, $0 \leq a_1 \leq a_2 \leq A$ and $y, v \in W(D)$,

$$(12) \quad \begin{aligned} & \int_{t_1}^{t_2} \int_{a_1}^{a_2} \langle y_t + y_a, v \rangle_V + \langle v_t + v_a, y \rangle_V \, da \, dt \\ &= \int_{a_1}^{a_2} (y(t_2, a), v(t_2, a))_H - (y(t_1, a), v(t_1, a))_H \, da \\ &+ \int_{t_1}^{t_2} (y(t, a_2), v(t, a_2))_H - (y(t, a_1), v(t, a_1))_H \, dt. \end{aligned}$$

3.2. Existence and uniqueness of solution. The aim of this section is to prove the existence of a solution of the state equation (5) in the following sense.

Definition 3.4. A function $y \in W(D)$ is a solution of (5) if

$$(13) \quad \begin{cases} \int_D \langle y_t + y_a, v \rangle_V \, da \, dt + \int_D \langle \mathcal{A}y, v \rangle_V \, da \, dt = \int_D \langle f, v \rangle_V \, da \, dt, & \forall v \in L^2(D; V), \\ y(0, a) = y_0(a), & \text{a.e. in } [0, A], \\ y(t, 0) = y_1(t), & \text{a.e. in } [0, T]. \end{cases}$$

The main idea to construct solutions of the state equation is to solve a family of parabolic differential equations over the characteristic curves where $t - a$ is constant. In the Appendix we recall some useful results for the standard parabolic case.

3.2.1. Existence. In this section we will construct a function given by the characteristic method and then we will prove that it is a solution of (5) in the sense of Definition 3.4. The idea is to solve an abstract parabolic differential equation over each characteristic curve, that depends only on time or age. We consider two regions D_1 and D_2 defined as

$$(14) \quad D_1 := \{(t, a) \in D : 0 \leq t \leq a \leq A\}, \quad D_2 := D \setminus D_1.$$

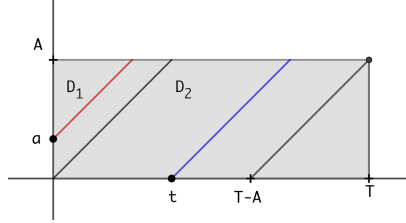


FIGURE 1. Partition of D

In the region D_1 we consider the curves that start in points with $t = 0$, i.e., for each $a \in [0, A]$ we define the characteristic curve as

$$(15) \quad D_1^a := \{(t, a + t) : 0 \leq t \leq A - a\}.$$

In the region D_2 we consider the curves defined as

$$(16) \quad D_2^t := \begin{cases} \{(t + a, a) : 0 \leq a \leq A\}, & \text{if } t \in (0, T - A], \\ \{(t + a, a) : 0 \leq a \leq T - t\}, & \text{if } t \in (T - A, T). \end{cases}$$

We start by solving the system (5) in the region D_1 . For $0 \leq a \leq A$, we denote by $Y^a(\cdot) \in W(0, A - a)$ the solution of the following system, which exists by Proposition A.1 for a.a. $a \in [0, A]$,

$$(17) \quad \begin{cases} \dot{Y}^a(t) + \mathcal{A}^a(t)Y^a(t) = f^a(t), & \text{in } L^2(0, A - a; V^*), \\ Y^a(0) = y_0(a), & \text{in } H, \end{cases}$$

where $\mathcal{A}^a(t) := \mathcal{A}(t, t + a)$, $f^a(t) := f(t, t + a)$, and, as usual, we denote $\dot{Y}^a(t)$ the time derivative of Y^a . By our assumptions, for almost all $a \in [0, A]$, we have $f^a \in L^2(0, A - a; V^*)$ and $y_0(a) \in H$, then Proposition A.1 applies. In particular, $Y^a \in W(0, A - a)$ satisfies

$$(18) \quad \int_0^{A-a} \langle \dot{Y}^a(t), v(t) \rangle_V \, dt + \int_0^{A-a} \langle \mathcal{A}^a(t)Y^a(t), v(t) \rangle_V \, dt = \int_0^{A-a} \langle f^a(t), v(t) \rangle_V \, dt,$$

for all $v \in L^2(0, A - a; V)$ and there exists $C > 0$ independent of a such that,

$$(19) \quad \|Y^a\|_{W(0, A-a)} \leq C [\|y_0(a)\|_H + \|f^a\|_{L^2(0, A-a; V^*)}].$$

Now, we define $D_1^* := \{(t, a) : 0 \leq a \leq A, 0 \leq t \leq A - a\}$. Based on [8, Lemma 1.8.2, p. 36] we obtain the following result:

Lemma 3.5. *The map $(t, a) \in D_1^* \mapsto Y^a(t)$ belongs to $L^2(D_1^*; V)$.*

Proof. We define the Banach space $E := H \times L^2(0, A; V^*)$. Given $(w, g) \in E$, by Proposition A.1 there exists a unique solution of the parabolic equation:

$$(20) \quad \begin{cases} \dot{z}(t) + \mathcal{A}^a(t)z(t) = g(t), & \text{in } L^2(0, A; V^*), \\ z(0) = w, & \text{in } H, \end{cases}$$

where the map \mathcal{A} is extended by zero out of D . Also, by Proposition A.1, the map $(w, g) \mapsto z$ is continuous and satisfies

$$(21) \quad \|z\|_{L^2(0, A; V)} \leq C[\|w\|_H + \|g\|_{L^2(0, A; V^*)}].$$

Since $y_0 \in L^2(0, A; H)$ and $f \in L^2(D; V^*)$, the map defined as

$$(22) \quad a \in [0, A] \mapsto (y_0(a), f^a),$$

belongs to $L^2(0, A; E)$ (we extend by zero the function f^a in $[A - a, A]$).

Therefore, by [8, Lemma 1.8.2, p. 36] the function $(t, a) \mapsto Y^a(t)$ belongs to $L^2(0, A; L^2(0, A; V))$, and by [8, Proposition 1.8.1, p. 28] the result follows. \square

Now, in region D_2 for $t \in (0, T - A]$ we consider the function \tilde{Y}^t solution of

$$(23) \quad \begin{cases} \dot{\tilde{Y}}^t(a) + \tilde{\mathcal{A}}^t(a)\tilde{Y}^t(a) = \tilde{f}^t(a), & \text{in } L^2(0, A; V^*), \\ \tilde{Y}^t(0) = y_1(t) & \text{in } H, \end{cases}$$

where $\tilde{\mathcal{A}}^t(a) := \mathcal{A}(t + a, a)$, $\tilde{f}^t(a) := f(t + a, a)$ and $\dot{\tilde{Y}}^t(a) = \frac{d}{da}\tilde{Y}^t(a)$. Again, by Proposition A.1, for a.a. $t \in (0, T - A]$ there exists $\tilde{Y}^t \in W(0, A)$ such that for all $v \in L^2(0, A; V)$,

$$(24) \quad \int_0^A \langle \dot{\tilde{Y}}^t(a), v(a) \rangle_V da + \int_0^A \langle \tilde{\mathcal{A}}^t(a)\tilde{Y}^t(a), v(a) \rangle_V da = \int_0^A \langle \tilde{f}^t(a), v(a) \rangle_V da,$$

and there exists $C > 0$, independent of t , such that

$$(25) \quad \|\tilde{Y}^t\|_{W(0, A)} \leq C[\|y_1(t)\|_H + \|\tilde{f}^t\|_{L^2(0, A; V^*)}].$$

We proceed analogously for $t \in (T - A, T]$ obtaining $\tilde{Y}^t \in W(0, T - t)$.

Now we define the function y as

$$(26) \quad y(t, a) := \begin{cases} Y^{a-t}(t) & (t, a) \in D_1, \\ \tilde{Y}^{t-a}(a) & (t, a) \in D_2. \end{cases}$$

We can extend the results of Lemma 3.5 to the region D_2 , and conclude that y belongs to $L^2(D; V)$. In addition, we have

$$(27) \quad \int_D \|y(t, a)\|_V^2 da dt = \int_{D_1} \|y(t, a)\|_V^2 da dt + \int_{D_2} \|y(t, a)\|_V^2 da dt.$$

By (26) and (19) we obtain

$$(28) \quad \begin{aligned} \int_{D_1} \|y(t, a)\|_V^2 da dt &= \int_0^A \int_0^{A-a} \|y(t, t+a)\|_V^2 dt da \\ &= \int_0^A \int_0^{A-a} \|Y^a(t)\|_V^2 dt da \\ &\leq C \int_0^A [\|y_0(a)\|_H^2 + \|f^a\|_{L^2(0, A-a; V^*)}^2] da. \end{aligned}$$

Analogously we can proceed in region D_2 and finally obtain

$$(29) \quad \|y\|_{L^2(D; V)}^2 \leq \bar{C} [\|y_0\|_{L^2(0, A; H)}^2 + \|y_1\|_{L^2(0, T; H)}^2 + \|f\|_{L^2(D; V^*)}^2].$$

In order to prove that y defined by (26) is a solution of the system (5) we need to prove that $y_t + y_a$ exists in the sense of (8), and belongs to $L^2(D; V^*)$. Let $\varphi \in C_0^\infty(D)$, then

$$(30) \quad \begin{aligned} \int_0^A \int_0^T y(t, a)(\varphi_t + \varphi_a)(t, a) dt da &= \int_{D_1} y(t, a)(\varphi_t + \varphi_a)(t, a) dt da \\ &\quad + \int_{D_2} y(t, a)(\varphi_t + \varphi_a)(t, a) dt da. \end{aligned}$$

Again, we start by D_1 . For all $a \in [0, A]$, the function $\varphi^a(t) := \varphi(t, t+a)$ belongs to $C_0^\infty(0, A-a)$, and $\dot{\varphi}^a(t) = \varphi_t(t, t+a) + \varphi_a(t, t+a)$. Since $Y^a \in W(0, A-a)$ we have

$$(31) \quad \int_0^{A-a} \dot{Y}^a(t) \varphi^a(t) dt = - \int_0^{A-a} Y^a(t) \dot{\varphi}^a(t) dt, \quad \text{in } V^*.$$

We can conclude that

$$(32) \quad \begin{aligned} \int_{D_1} y(t, a) (\varphi_t + \varphi_a)(t, a) dt da &= \int_0^A \int_0^{A-a} y(t, t+a) (\varphi_t + \varphi_a)(t, t+a) dt da \\ &= \int_0^A \int_0^{A-a} Y^a(t) \dot{\varphi}^a(t) dt da \\ &= - \int_0^A \int_0^{A-a} \dot{Y}^a(t) \varphi^a(t) dt da \\ &= \int_{D_1} \dot{Y}^{a-t}(t) \varphi(t, a) da dt. \end{aligned}$$

We can proceed analogously in D_2 and conclude that $y_t + y_a$ exists in a weak sense in D and satisfies

$$(33) \quad (y_t + y_a)(t, a) = \begin{cases} \dot{Y}^{a-t}(t), & \text{if } (t, a) \in D_1, \\ \dot{Y}^{t-a}(a), & \text{if } (t, a) \in D_2. \end{cases}$$

In this case, the well posedness of $y_t + y_a$ is also guaranteed by Lemma 3.5, obtaining $y_t + y_a \in L^2(D; V^*)$. We can conclude that the function defined by (26) belongs to $W(D)$.

Finally, by (18), (24), (26) and (33), integrating over the characteristic curves in D , we can conclude that the function y is a solution of (5) in the sense of Definition 3.4.

3.2.2. Uniqueness. We start by giving an *a priori* estimate.

Proposition 3.6. *Let $y \in W(D)$ be a solution of (5), then there exists $C > 0$ such that*

$$(34) \quad \sup_{t \in [0, T]} \|y(t, \cdot)\|_{L^2(0, A; H)}^2 \leq C \left[\|y_0\|_{L^2(0, A; H)}^2 + \|y_1\|_{L^2(0, T; H)}^2 + \|f\|_{L^2(D; V^*)}^2 \right],$$

$$(35) \quad \sup_{a \in [0, A]} \|y(\cdot, a)\|_{L^2(0, T; H)}^2 \leq C \left[\|y_0\|_{L^2(0, A; H)}^2 + \|y_1\|_{L^2(0, T; H)}^2 + \|f\|_{L^2(D; V^*)}^2 \right],$$

and

$$(36) \quad \|y\|_{W(D)}^2 \leq C \left[\|y_0\|_{L^2(0, A; H)}^2 + \|y_1\|_{L^2(0, T; H)}^2 + \|f\|_{L^2(D; V^*)}^2 \right].$$

Proof. By Lemma 3.3, taking $y = v$ in (13), for all $\bar{t} \in [0, T]$ and $\bar{a} \in [0, A]$ we have,

$$(37) \quad \begin{aligned} \int_0^{\bar{t}} \int_0^{\bar{a}} \langle y_t + y_a, y \rangle_V da dt &= \frac{1}{2} \int_0^{\bar{a}} (\|y(\bar{t}, a)\|_H^2 - \|y_0(a)\|_H^2) da \\ &\quad + \frac{1}{2} \int_0^{\bar{t}} (\|y(t, \bar{a})\|_H^2 - \|y_1(t)\|_H^2) dt. \end{aligned}$$

Since $y \in L^2(D; V)$, by (13) we obtain

$$(38) \quad \begin{aligned} \frac{1}{2} \int_0^{\bar{a}} \|y(\bar{t}, a)\|_H^2 da + \frac{1}{2} \int_0^{\bar{t}} \|y(t, \bar{a})\|_H^2 dt + \int_0^{\bar{t}} \int_0^{\bar{a}} \langle \mathcal{A}(t, a) y(t, a), y(t, a) \rangle_V da dt \\ = \int_0^{\bar{t}} \int_0^{\bar{a}} \langle f, y \rangle_V da dt + \frac{1}{2} \int_0^{\bar{a}} \|y_0\|_H^2 da + \frac{1}{2} \int_0^{\bar{t}} \|y_1\|_H^2 dt. \end{aligned}$$

We take $\bar{a} = A$ in (38), then by the Cauchy-Schwarz and Young inequalities, and the semicoercivity assumption for $\mathcal{A}(t, a)$ we obtain for all $\varepsilon > 0$,

$$(39) \quad \begin{aligned} \int_0^A \|y(\bar{t}, a)\|_H^2 da &\leq -2 \int_0^{\bar{t}} \int_0^A \langle \mathcal{A}(t, a) y(t, a), y(t, a) \rangle_V da dt + \|y_0\|_{L^2(0, A; H)}^2 \\ &\quad + \|y_1\|_{L^2(0, T; H)}^2 + 2 \int_0^{\bar{t}} \int_0^A \|f\|_{V^*} \|y\|_V da dt \\ &\leq 2\lambda \int_0^{\bar{t}} \int_0^A \|y\|_H^2 da dt - 2\alpha \int_0^{\bar{t}} \int_0^A \|y\|_V^2 da dt + \|y_0\|_{L^2(0, A; H)}^2 \\ &\quad + \|y_1\|_{L^2(0, T; H)}^2 + \frac{1}{\varepsilon} \|f\|_{L^2(D; V^*)}^2 + \varepsilon \int_0^{\bar{t}} \int_0^A \|y\|_V^2 da dt \\ &= 2\lambda \int_0^{\bar{t}} \int_0^A \|y\|_H^2 da dt + (\varepsilon - 2\alpha) \int_0^{\bar{t}} \int_0^A \|y\|_V^2 da dt \\ &\quad + \|y_0\|_{L^2(0, A; H)}^2 + \|y_1\|_{L^2(0, T; H)}^2 + \frac{1}{\varepsilon} \|f\|_{L^2(D; V^*)}^2. \end{aligned}$$

Taking $\varepsilon = \alpha$, and defining the function $\beta(s) := \int_0^A \|y(s, a)\|_H^2 da$, we have

$$(40) \quad \beta(\bar{t}) \leq 2\lambda \int_0^{\bar{t}} \beta(t) dt + \|y_0\|_{L^2(0,A;H)}^2 + \|y_1\|_{L^2(0,T;H)}^2 + \frac{1}{\alpha} \|f\|_{L^2(D;V^*)}^2.$$

By Grönwall's Lemma, there exists $C > 0$ such that

$$(41) \quad \sup_{s \in [0, T]} \int_0^A \|y(s, a)\|_H^2 da \leq C e^{2\lambda T} \left[\|y_0\|_{L^2(0,A;H)}^2 + \|y_1\|_{L^2(0,T;H)}^2 + \|f\|_{L^2(D;V^*)}^2 \right].$$

Analogously we obtain

$$(42) \quad \sup_{b \in [0, A]} \int_0^T \|y(t, b)\|_H^2 dt \leq C e^{2\lambda A} \left[\|y_0\|_{L^2(0,A;H)}^2 + \|y_1\|_{L^2(0,T;H)}^2 + \|f\|_{L^2(D;V^*)}^2 \right].$$

We conclude that (34) and (35) hold.

Then, by (34) or (35) there exists $C > 0$ such that

$$(43) \quad \|y\|_{L^2(D;H)}^2 \leq C \left[\|y_0\|_{L^2(0,A;H)}^2 + \|y_1\|_{L^2(0,T;H)}^2 + \|f\|_{L^2(D;V^*)}^2 \right].$$

Now, going back to (39), and taking $\bar{t} = T$, we have

$$(44) \quad \alpha \int_D \|y\|_V^2 da dt \leq 2\lambda \int_D \|y\|_H^2 da dt + \|y_0\|_{L^2(0,A;H)}^2 + \|y_1\|_{L^2(0,T;H)}^2 + \frac{1}{\alpha} \|f\|_{L^2(D;V^*)}^2.$$

And by (43) and the previous inequality there exists $C > 0$ such that

$$(45) \quad \|y\|_{L^2(D;V)}^2 \leq C \left[\|y_0\|_{L^2(0,A;H)}^2 + \|y_1\|_{L^2(0,T;H)}^2 + \|f\|_{L^2(D;V^*)}^2 \right].$$

Now, since y is solution of (5), we have

$$(46) \quad \|y_t + y_a\|_{L^2(D;V^*)} \leq \|\mathcal{A}y\|_{L^2(D;V^*)} + \|f\|_{L^2(D;V^*)}.$$

By (H1) and (45), there exists $C > 0$ such that

$$(47) \quad \|y\|_{W(D)}^2 \leq C \left[\|y_0\|_{L^2(0,A;H)}^2 + \|y_1\|_{L^2(0,T;H)}^2 + \|f\|_{L^2(D;V^*)}^2 \right].$$

□

If we assume that there exist two solutions y and z of (13), then their difference $y - z$ is solution of the same equation with zero r.h.s. and zero initial conditions, by the previous proposition we conclude $y = z$ in $W(D)$. To summarize, we have

Theorem 3.7. *Under assumptions (H1) and (H2), there exists a unique solution of (5) in $W(D)$ which is given by (26).*

Remark 3.8. By the previous theorem we conclude that the map

$$(48) \quad y \mapsto (y_t + y_a + \mathcal{A}y, y(0, \cdot), y(\cdot, 0)),$$

defines an isomorphism between $W(D)$ and $L^2(D; V^*) \times L^2(0, A; H) \times L^2(0, T; H)$.

Furthermore, we deduce that for any $z \in W(D)$, the function defined as $z^a(t) := z(t, t+a)$ belongs to $W(0, A-a)$ for a.a. $a \in [0, A]$ and $\tilde{z}^t(a) := z(t+a, a)$ belongs to $W(0, A \wedge (T-t))$ for a.a. $t \in [0, T]$.

3.2.3. Second parabolic estimates. The estimates studied in Proposition 3.6 are usually called the first parabolic estimates. Now, under additional assumptions we show the second parabolic estimate for the age-structured case. We make the following assumptions: $\mathcal{A}(t, a) = \mathcal{A}_0(t, a) + \mathcal{A}_1(t, a)$ where

(H3) $\mathcal{A}_0(t, a) : V \rightarrow V^*$ is symmetric and uniformly continuously differentiable over almost all characteristic curves, i.e., defining for a.a. $a \in [0, A]$, the function $\mathcal{A}^a(t) := \mathcal{A}(t, t+a)$ and for a.a. $t \in [0, T]$, the function $\tilde{\mathcal{A}}^t(a) := \mathcal{A}(t+a, a)$, there exists a modulus of continuity for \mathcal{A}^a , $\dot{\mathcal{A}}^a$, $\tilde{\mathcal{A}}^t$ and $\dot{\tilde{\mathcal{A}}}^t$, independent of t and a . In addition, there exists $\alpha_0 > 0$ such that $\langle \mathcal{A}_0(t, a)v, v \rangle \geq \alpha_0 \|v\|_V^2$, for all $v \in V$ and a.a. $(t, a) \in D$.

(H4) $\mathcal{A}_1(t, a) : V \rightarrow V^*$ is measurable with range in H and there exists $c_1 > 0$ such that $\|\mathcal{A}_1(t, a)v\|_H \leq c_1 \|v\|_V$, for all $v \in V$ and a.a. $(t, a) \in D$.

Theorem 3.9. Assume $f \in L^2(D; H)$, $y_0 \in L^2(0, A; V)$, $y_1 \in L^2(0, T; V)$ and the assumptions (H1)-(H4) hold. Then, the solution y of the system (5) satisfies

$$(49) \quad y \in L^\infty(0, T; L^2(0, A; V)) \cap L^\infty(0, A; L^2(0, T; V)), \quad y_t + y_a \in L^2(D; H),$$

and there exists $C > 0$ such that

$$(50) \quad \begin{aligned} & \|y\|_{L^\infty(0, T; L^2(0, A; V))} + \|y\|_{L^\infty(0, A; L^2(0, T; V))} + \|y_t + y_a\|_{L^2(D; H)} \\ & \leq C [\|y_0\|_{L^2(0, A; V)} + \|y_1\|_{L^2(0, T; V)} + \|f\|_{L^2(D; H)}]. \end{aligned}$$

Proof. We know that the solution y is defined over the characteristic curves. In particular, for almost all characteristic Remark A.3 applies.

Consider $\bar{a} \in [0, A]$. We want to prove that $y(\cdot, \bar{a}) \in L^2(0, T; V)$. For every $t \in [0, T]$, (t, \bar{a}) belongs to a characteristic curve. In particular, if $t \leq \bar{a}$, then $(t, \bar{a}) \in D_1$, and the value of $y(t, \bar{a})$ is given by the function $Y^{\bar{a}-t}(t)$. By Remark A.3, we have

$$(51) \quad \|y(t, \bar{a})\|_V \leq C [\|y_0(\bar{a} - t)\|_V + \|f^{\bar{a}-t}\|_{L^2(0, A-(\bar{a}-t); H)}], \quad \text{a.e. } t \in [0, \bar{a}],$$

where $C > 0$ is independent of y_0 and f .

On the other hand, if $T - A \geq t > \bar{a}$, the value of $y(t, \bar{a})$ is given by $\tilde{Y}^{t-\bar{a}}(\bar{a})$. Again, by Remark A.3 for a.e. $t \in [\bar{a}, T]$ we obtain

$$(52) \quad \|y(t, \bar{a})\|_V \leq C [\|y_1(t - \bar{a})\|_V + \|\tilde{f}^{t-\bar{a}}\|_{L^2(0, A; H)}].$$

When $T - A < t \leq T$, we obtain similar estimates, where the last term in the r.h.s. of the previous inequality has to be understood in $L^2(0, T - t; H)$.

By (51) and (52) there exists $\bar{C} > 0$, independent of \bar{a} , such that

$$(53) \quad \int_0^T \|y(t, \bar{a})\|_V^2 dt \leq \bar{C} [\|y_0\|_{L^2(0, A; V)}^2 + \|y_1\|_{L^2(0, T; V)}^2 + \|f\|_{L^2(D; H)}^2].$$

Analogously we can prove that for all $\bar{t} \in [0, T]$, $y(\bar{t}, \cdot) \in L^2(0, A; V)$ and we obtain

$$(54) \quad \begin{aligned} & \|y\|_{L^\infty(0, T; L^2(0, A; V))} + \|y\|_{L^\infty(0, A; L^2(0, T; V))} \\ & \leq C [\|y_0\|_{L^2(0, A; V)} + \|y_1\|_{L^2(0, T; V)} + \|f\|_{L^2(D; H)}]. \end{aligned}$$

Following similar arguments, by (32), for almost all $(t, a) \in D$ the value of $y_t + y_a$ is given by \dot{Y}^{a-t} in D_1 or by $\dot{\tilde{Y}}^{t-a}$ in D_2 . Then, by Theorem A.2 we obtain $(y_t + y_a)(t, a) \in H$ a.e. and also there exists $C > 0$ such that

$$(55) \quad \|y_t + y_a\|_{L^2(D; H)} \leq C [\|y_0\|_{L^2(0, A; V)} + \|y_1\|_{L^2(0, T; V)} + \|f\|_{L^2(D; H)}].$$

By (54) and (55) the result follows. \square

3.3. Birth process depending on a fecundity rate. It is natural in many applications to consider that the birth process $y(\cdot, 0)$ depends on a fecundity rate. We notice that we can also prove the existence and uniqueness of solution for the following system that takes this into account,

$$(56) \quad \begin{cases} (y_t + y_a)(t, a) + \mathcal{A}(t, a)y(t, a) = f(t, a) & \text{in } L^2(D; V^*), \\ y(0, a) = y_0(a) & \text{in } L^2(0, A; H), \\ y(t, 0) = \int_0^A c(t, a)y(t, a)da & \text{in } L^2(0, T; H). \end{cases}$$

In this case $c : D \rightarrow \mathbb{R}^+$ is a measurable function that represents the fecundity rate, and we assume that there exists $\bar{C} > 0$ such that

$$(57) \quad \sup_{t \in [0, T]} \int_0^A c(a, t)^2 da \leq \bar{C}.$$

Let $\gamma > 0$, and define $w(t, a) := e^{-\gamma t}y(t, a)$. It is easy to observe that if y satisfies the first equation in (56), then w satisfies

$$(58) \quad (w_t + w_a)(t, a) + \mathcal{A}(t, a)w(t, a) + \gamma w = \bar{f}(t, a) \quad \text{in } L^2(D; V^*),$$

where $\bar{f}(t, a) := e^{-\gamma t} f(t, a)$. So, in the sequel we consider the above dynamics. Now, we define the map $\mathcal{T} : L^2(D; H) \rightarrow L^2(D; H)$, such that $\mathcal{T}z = w$, where w solves the system

$$(59) \quad \begin{cases} (w_t + w_a)(t, a) + \mathcal{A}(t, a)w(t, a) + \gamma w = \bar{f}(t, a) & \text{in } L^2(D; V^*), \\ w(0, a) = w_0(a) & \text{in } L^2(0, A; H), \\ w(t, 0) = \int_0^A c(t, a)z(t, a)da & \text{in } L^2(0, T; H). \end{cases}$$

We observe that by the Cauchy-Schwarz inequality and our assumptions we have

$$(60) \quad \int_0^T \left\| \int_0^A c(t, a)z(t, a)da \right\|_H^2 dt \leq \int_0^T \left(\int_0^A c(t, a)^2 da \right) \left(\int_0^A \|z(t, a)\|_H^2 da \right) dt \leq \bar{C} \|z\|_{L^2(D; H)}^2.$$

Then, the map $t \mapsto \int_0^A c(t, a)z(t, a)da$ belongs to $L^2(0, T; H)$ and by Theorem 3.7, there exists a unique solution of (59). We conclude that the map \mathcal{T} is well-posed. Now we want to prove that it is a contractive map in $L^2(D; H)$. Let $z_1, z_2 \in L^2(D; H)$, we denote $\bar{z} := z_1 - z_2$ and $\bar{w} := \mathcal{T}z_1 - \mathcal{T}z_2$, we have that \bar{w} solves the system (59) with $f \equiv 0$, $w_0 \equiv 0$ and \bar{z} instead of z .

By Lemma 3.3 and assumption (H2), we obtain

$$(61) \quad \begin{aligned} 0 &= \int_D [\langle \bar{w}_t + \bar{w}_a, \bar{w} \rangle_V + \langle \mathcal{A}\bar{w}, \bar{w} \rangle_V + \gamma \langle \bar{w}, \bar{w} \rangle_V] dadt \\ &= \frac{1}{2} \int_0^A [\|\bar{w}(T, a)\|_H^2 - \|\bar{w}(0, a)\|_H^2] da + \frac{1}{2} \int_0^T [\|\bar{w}(t, A)\|_H^2 - \|\bar{w}(t, 0)\|_H^2] dt \\ &\quad + \int_D [\langle \mathcal{A}\bar{w}, \bar{w} \rangle_V + \gamma \langle \bar{w}, \bar{w} \rangle_V] dadt, \\ &\geq -\frac{1}{2} \int_0^T \|\bar{w}(t, 0)\|_H^2 dt + \int_D [\alpha \|\bar{w}\|_V^2 - \lambda \|\bar{w}\|_H^2 + \gamma \|\bar{w}\|_H^2] dadt. \end{aligned}$$

We deduce with (60) that

$$(62) \quad (\gamma - \lambda) \|\bar{w}\|_{L^2(D; H)}^2 \leq \frac{1}{2} \int_0^T \|w(t, 0)\|_H^2 dt \leq \frac{\bar{C}}{2} \|\bar{z}\|_{L^2(D; H)}^2.$$

We can choose $\gamma > 0$ large enough in order to obtain that \mathcal{T} is a contraction, and therefore we have a unique fixed point. In addition, by Theorem 3.7, we have that $\mathcal{T}(L^2(D; H)) \subset W(D)$, then, the system (56) has a unique solution in $W(D)$.

If in addition we assume that \mathcal{A} satisfies (H1)-(H4), $f \in L^2(D; H)$ and $y_0 \in L^2(0, A; V)$, since the solution of (56) belongs to $W(D)$, we have $y(\cdot, 0) \in L^2(D; V)$, and then Theorem 3.9 applies.

4. OPTIMALITY CONDITION

The aim of this section is provide optimality conditions for the optimal control problem described in Section 2. In order to analyze the differentiability of the cost function, we study the regularity of the map $u \mapsto y[u]$, by the Implicit Function Theorem. We also set the space for the state constraints and the associated multipliers.

We recall the optimal control problem that we consider

$$(63) \quad \min J(u, y) := \int_0^T \int_0^A \ell(u, y) dadt + \int_0^A h_0(y(T, a)) da + \int_0^T h_1(y(t, A)) dt$$

s.t.

$$(64) \quad [p] \quad (y_t + y_a)(t, a) + \mathcal{A}(t, a, u(t, a))y(t, a) = f(t, a), \quad \text{in } L^2(D; V^*),$$

$$(65) \quad [\psi_0] \quad y(0, a) - y_0(a) = 0, \quad \text{in } L^2(0, A; H),$$

$$(66) \quad [\psi_1] \quad y(t, 0) - y_1(t) = 0, \quad \text{in } L^2(0, T; H),$$

$$(67) \quad [\mu_i] \quad \int_0^A (\xi_i(a), y(t, a))_H da - M_i \leq 0, \quad t \in [0, T], \quad i = 1, \dots, r,$$

and

$$(68) \quad u \in K_{\mathcal{U}},$$

where $K_{\mathcal{U}}$ is a nonempty, closed and convex subset of \mathcal{U} . In the brackets in the column on the left we write the multipliers associated to each constraint, that will be used in the definition of the Lagrangian function.

In order to apply the results of Section 3 to the state equation (64)-(66), along this section we consider the following assumptions:

- (H5) The map $\mathcal{A} : D \times U \rightarrow L(V, V^*)$ is measurable w.r.t. (t, a) for all $u \in U$, and is C^1 w.r.t. u for a.a. $(t, a) \in D$.
- (H6) For \mathcal{A} and \mathcal{A}_u there exists a modulus of continuity w.r.t. u , on bounded sets of U , which is independent of (t, a) .
- (H7) The assumptions (H1) and (H2) are satisfied by $\mathcal{A}(t, a, u)$, for a.a. $(t, a) \in D$, with uniform coefficients C in (H1) and α and λ in (H2), over bounded sets of U .
- (H8) $\xi_i \in W(0, A)$, for $i = 1, \dots, r$.

Remark 4.1. By the assumption (H5), for any $u \in \mathcal{U}$ we can define the following measurable map,

$$(69) \quad \begin{aligned} \mathcal{A}[u] : \quad D &\rightarrow L(V, V^*) \\ (t, a) &\mapsto \mathcal{A}(t, a, u(t, a)). \end{aligned}$$

By assumptions (H5)-(H7), we can define the Nemitskii map $u \in \mathcal{U} \mapsto \mathcal{A}[u] \in L^\infty(D; L(V, V^*))$, which is C^1 and satisfies

$$(70) \quad (D_u \mathcal{A}[u]v)(t, a) = D_u \mathcal{A}(t, a, u(t, a))v(t, a).$$

In fact, by (H7) since $u \in \mathcal{U}$, we can deduce that $\mathcal{A}[u] \in L^\infty(D; L(V, V^*))$. By (H6),

$$(71) \quad \|\mathcal{A}[u + v] - \mathcal{A}[u] - D_u \mathcal{A}(u)v\|_{L^\infty(D; L(V, V^*))} = o(\|v\|_{\mathcal{U}}).$$

4.1. Preliminary results.

Lemma 4.2. *Under the above assumptions, the map $y[u, y_0, y_1, f]$ is C^1 , where $y[u, y_0, y_1, f] \in W(D)$ solves the system (64)-(66) with $u \in \mathcal{U}$, $y_0 \in L^2(0, A; H)$, $y_1 \in L^2(0, T; H)$ and $f \in L^2(D; V^*)$.*

Proof. We define the function $\Psi : W(D) \times \mathcal{U} \times L^2(0, A; H) \times L^2(0, T; H) \times L^2(D; V^*) \rightarrow L^2(D; V^*) \times L^2(0, A; H) \times L^2(0, T; H)$ as

$$(72) \quad \Psi(y, u, y_0, y_1, f) := \begin{pmatrix} f - (y_t + y_a + \mathcal{A}(u)y) \\ y_0 - y(0, \cdot) \\ y_1 - y(\cdot, 0) \end{pmatrix}.$$

By Remark 4.1, we obtain that the map Ψ is C^1 w.r.t. all the variables. In addition, by (H7) and the results of Section 3, for all $(g, x_0, x_1) \in L^2(D; V^*) \times L^2(0, A; H) \times L^2(0, T; H)$, there exists a unique solution of

$$(73) \quad \begin{cases} z_t + z_a + \mathcal{A}(u)z = g & \text{in } L^2(D; V^*), \\ z(0, \cdot) = x_0, & \text{in } L^2(0, A; H), \\ z(\cdot, 0) = x_1, & \text{in } L^2(0, T; H). \end{cases}$$

And that implies that $D_y \Psi$ is invertible. Then, by the Implicit Function Theorem, the result follows. \square

In order to define the Lagrangian of the problem, we need to specify in which space we consider the state constraints. The following lemma proves that we can take the space $C([0, T])$ of continuous functions on $[0, T]$.

Lemma 4.3. *Let $\xi \in W(0, A)$ and $y \in W(D)$, then $\int_0^A (\xi(a), y(t, a))_H da$ is continuous with respect to $t \in [0, T]$.*

Proof. By Remark 3.8, we can define $(f, y_0, y_1) \in L^2(D; V^*) \times L^2(0, A; H) \times L^2(0, T; H)$ as

$$(74) \quad f(t, a) := (y_t + y_a)(t, a) + \mathcal{A}(t, a)y(t, a), \quad y_0(a) := y(0, a), \quad \text{and} \quad y_1(t) := y(t, 0),$$

where \mathcal{A} satisfies (H1)-(H2) (for instance we can take $\langle \mathcal{A}y, v \rangle := (y, v)_V$, the scalar product in V).

Since $y \in W(D)$ we know that for all $t \in [0, T]$, $y(t, \cdot) \in L^2(0, A; H)$. Let $\varepsilon > 0$, and $t \in [0, T - \varepsilon]$ then

$$(75) \quad \begin{aligned} \left| \int_0^A (\xi(a), y(t + \varepsilon, a) - y(t, a))_H da \right| &\leq \left| \int_0^\varepsilon (\xi(a), y(t + \varepsilon, a) - y(t, a))_H da \right| \\ &+ \left| \int_\varepsilon^{A-\varepsilon} (\xi(a), y(t + \varepsilon, a) - y(t, a))_H da \right| \\ &+ \left| \int_{A-\varepsilon}^A (\xi(a), y(t + \varepsilon, a) - y(t, a))_H da \right|. \end{aligned}$$

For the first term, by the Cauchy-Schwarz inequality and Proposition 3.6, we have

$$(76) \quad \begin{aligned} & \left(\int_0^\varepsilon (\xi(a), y(t + \varepsilon, a) - y(t, a))_H da \right)^2 \\ & \leq \int_0^\varepsilon \|\xi(a)\|_H^2 da \int_0^A \|y(t + \varepsilon, a) - y(t, a)\|_H^2 da \\ & \leq 2C [\|y_0\|_{L^2}^2 + \|y_1\|_{L^2}^2 + \|f\|_{L^2}^2] \int_0^\varepsilon \|\xi(a)\|_H^2 da. \end{aligned}$$

And by the Dominated Convergence Theorem we have

$$(77) \quad \int_0^\varepsilon \|\xi(a)\|_H^2 da \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0.$$

For the last term in (75), similar estimates holds. Now, for the second term we obtain

$$(78) \quad \begin{aligned} & \left| \int_\varepsilon^{A-\varepsilon} (\xi(a), y(t + \varepsilon, a) - y(t, a))_H da \right| \\ & \leq \left| \int_\varepsilon^{A-\varepsilon} (\xi(a), y(t + \varepsilon, a) - y(t + \varepsilon, a + \varepsilon))_H da \right| \\ & \quad + \left| \int_\varepsilon^{A-\varepsilon} (\xi(a), y(t + \varepsilon, a + \varepsilon) - y(t, a))_H da \right|. \end{aligned}$$

We have

$$(79) \quad \begin{aligned} & \int_\varepsilon^{A-\varepsilon} (\xi(a), y(t + \varepsilon, a) - y(t + \varepsilon, a + \varepsilon))_H da = \int_\varepsilon^{2\varepsilon} (\xi(a), y(t + \varepsilon, a))_H da \\ & \quad - \int_{A-\varepsilon}^A (\xi(a - \varepsilon), y(t + \varepsilon, a))_H da + \int_{2\varepsilon}^{A-\varepsilon} (\xi(a) - \xi(a - \varepsilon), y(t + \varepsilon, a))_H da. \end{aligned}$$

The first two terms can be estimated as before and for the last one, by the Cauchy-Schwarz inequality we obtain

$$(80) \quad \begin{aligned} & \left| \int_{2\varepsilon}^{A-\varepsilon} (\xi(a) - \xi(a - \varepsilon), y(t + \varepsilon, a))_H da \right|^2 \\ & \leq \|y(t + \varepsilon, \cdot)\|_{L^2(0, A; H)}^2 \int_{2\varepsilon}^{A-\varepsilon} \|\xi(a) - \xi(a - \varepsilon)\|_H^2 da. \end{aligned}$$

Since $\xi \in W(0, A)$, the integral in the r.h.s. converges to zero and by Proposition 3.6, the first term is bounded, independently of ε , then the above term tends to zero when $\varepsilon \rightarrow 0$.

Now, by the Dominated Convergence Theorem we prove that the last term in (78) converges to zero. In fact, we have

$$(81) \quad \begin{aligned} & \left| \int_\varepsilon^{A-\varepsilon} (\xi(a), y(t + \varepsilon, a + \varepsilon) - y(t, a))_H da \right|^2 \\ & \leq \|\xi\|_{L^2(0, A; H)}^2 \int_\varepsilon^{A-\varepsilon} \|y(t + \varepsilon, a + \varepsilon) - y(t, a)\|_H^2 da. \end{aligned}$$

And for all $a \in [\varepsilon, A - \varepsilon]$, the points (t, a) and $(t + \varepsilon, a + \varepsilon)$ belong to the same characteristic curve. Since the solutions of the parabolic equations along the characteristic curves are continuous with images in H , we can conclude that, for a.a. $a \in [\varepsilon, A - \varepsilon]$, we have

$$(82) \quad \|y(t + \varepsilon, a + \varepsilon) - y(t, a)\|_H^2 \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0.$$

In addition, by the first parabolic estimate (Proposition A.1) we have (assuming $f \equiv 0$ out of D),

$$\begin{aligned}
 & \|y(t + \varepsilon, a + \varepsilon) - y(t, a)\|_H^2 \\
 & \leq 2 [\|y(t + \varepsilon, a + \varepsilon)\|_H^2 + \|y(t, a)\|_H^2] \\
 (83) \quad & \leq C \left\{ \chi_{[a > t]} \left[\|y_0(a - t)\|_H^2 + \int_0^A \|f(s, a - t + s)\|_{V^*}^2 ds \right] \right. \\
 & \quad \left. + \chi_{[a \leq t]} \left[\|y_1(t - a)\|_H^2 + \int_0^A \|f(t - a + s, s)\|_{V^*}^2 ds \right] \right\}.
 \end{aligned}$$

By our assumptions, the r.h.s. belongs to $L^1(0, A; H)$. We can proceed analogously for $\varepsilon < 0$ and $t \in [-\varepsilon, T]$, which completes the proof. \square

Since the space of the state constraints is $C([0, T])$, the associated multiplier will belong to the dual space, which is the set of Borel measure $M(0, T)$. It is known that every element in this space can be identified with $d\eta$ for some $\eta \in BV_T(0, T)$, where $BV(0, T)$ denote the space of bounded variation functions on $[0, T]$ and

$$(84) \quad BV_T(0, T) := \{\eta \in BV(0, T) : \eta(T) = 0\}.$$

For any $\eta \in BV_T(0, T)$ and $\varphi \in C([0, T])$ we have

$$(85) \quad \langle \eta, \varphi \rangle_{C([0, T])} := \int_0^T \varphi(t) d\eta(t).$$

In what follows we need the following result.

Lemma 4.4. *Let $\xi \in W(0, A)$, $\mu \in BV(0, T)$, and $z \in W(D)$. Then, the following integration by parts formula holds*

$$\begin{aligned}
 & \int_0^T \int_0^A \langle z_t + z_a, \xi(a) \rangle_V \mu(t) da dt + \int_0^T \int_0^A \langle \dot{\xi}(a), z \rangle_V \mu(t) da dt \\
 (86) \quad & = \int_0^A (z(T, a), \xi(a)) \mu(T) da - \int_0^A (z(0, a), \xi(a)) \mu(0) da \\
 & \quad + \int_0^T (z(t, A), \xi(A)) \mu(t) dt - \int_0^T (z(t, 0), \xi(0)) \mu(t) dt - \int_0^T \int_0^A (\xi(a), z)_H da d\mu(t).
 \end{aligned}$$

Proof. Defining $z^a(t) := z(t, t + a)$ and $\tilde{z}^t(a) := z(t + a, a)$, and integrating over the characteristic curves, we obtain

$$\begin{aligned}
 \int_D \langle (z_t + z_a)(t, a), \xi(a) \rangle_V \mu(t) da dt &= \int_0^A \int_0^{A-a} \langle \dot{z}^a(t), \xi(t + a) \rangle_V \mu(t) dt da \\
 (87) \quad & + \int_0^{T-A} \int_0^A \langle \dot{\tilde{z}}^t(a), \xi(a) \rangle_V \mu(t + a) da dt \\
 & + \int_{T-A}^T \int_0^{T-t} \langle \dot{\tilde{z}}^t(a), \xi(a) \rangle_V \mu(t + a) da dt.
 \end{aligned}$$

By Remark 3.8, $z^a \in W(0, A - a) \subset C(0, A - a; H)$. Then, for a.a. $a \in [0, A]$ the map $t \mapsto (z^a(t), \xi(t + a))_H$ is continuous and also we have

$$(88) \quad \langle \dot{z}^a(t), \xi(t + a) \rangle_V + \langle \dot{\xi}(t + a), z^a(t) \rangle_V \in L^1(0, A - a).$$

Since $\mu \in BV_T(0, T)$, by [4, Lemma 3.6] over each characteristic curve holds

$$\begin{aligned}
 \int_0^{A-a} \langle \dot{z}^a(t), \xi(t + a) \rangle_V \mu(t) dt &= - \int_0^{A-a} \langle \dot{\xi}(t + a), z^a(t) \rangle_V \mu(t) dt \\
 (89) \quad & - \int_0^{A-a} (z^a(t), \xi(t + a))_H d\mu(t) \\
 & + (z^a(A - a), \xi(A))_H \mu(A - a) \\
 & - (z^a(0), \xi(a))_H \mu(0).
 \end{aligned}$$

Then, since $z^a(A - a) = z(A - a, A)$ and $z^a(0) = z(0, a)$ we have

$$\begin{aligned}
 \int_0^A \int_0^{A-a} \langle \dot{z}^a(t), \xi(t + a) \rangle_V \mu(t) dt da &= - \int_0^A \int_0^{A-a} \langle \dot{\xi}(t + a), z^a(t) \rangle_V \mu(t) dt da \\
 (90) \quad & - \int_0^A \int_0^{A-a} (z^a(t), \xi(t + a))_H d\mu(t) da \\
 & + \int_0^A (z(t, A), \xi(A))_H \mu(t) dt \\
 & - \int_0^A (z(0, a), \xi(a))_H \mu(0) da.
 \end{aligned}$$

In a similar way, since $\tilde{z}^t(a) = z(t + A, A)$ and $\tilde{z}^t(0) = z(t, 0)$ we obtain

$$(91) \quad \begin{aligned} \int_0^{T-A} \int_0^A \langle \dot{\tilde{z}}^t(a), \xi(a) \rangle_V \mu(t+a) da dt &= - \int_0^{T-A} \int_0^A \langle \dot{\xi}(a), \tilde{z}^t(a) \rangle_V \mu(t+a) da dt \\ &\quad - \int_0^{T-A} \int_0^A (\tilde{z}^t(a), \xi(a))_H d\mu(t+a) dt \\ &\quad + \int_A^T (z(t, A), \xi(A))_H \mu(t) dt \\ &\quad - \int_0^{T-A} (z(t, 0), \xi(0))_H \mu(0) dt. \end{aligned}$$

And finally, since $\tilde{z}^t(T-t) = z(T, T-t)$ we have

$$(92) \quad \begin{aligned} \int_{T-A}^T \int_0^{T-t} \langle \dot{\tilde{z}}^t(a), \xi(a) \rangle_V \mu(t+a) da dt &= - \int_{T-A}^T \int_0^{T-t} \langle \dot{\xi}(a), \tilde{z}^t(a) \rangle_V \mu(t+a) da dt \\ &\quad - \int_{T-A}^T \int_0^{T-t} (\tilde{z}^t(a), \xi(a))_H d\mu(t+a) dt \\ &\quad + \int_0^A (z(T, a), \xi(a))_H \mu(T) da \\ &\quad - \int_{T-A}^T (z(t, 0), \xi(0))_H \mu(0) dt. \end{aligned}$$

By adding the last three equations the result follows. \square

4.2. Optimality conditions. In order to analyze the differentiability of the cost function, we add the following assumptions:

- (H9) The function $\ell : D \times U \times H \rightarrow \mathbb{R}$ is measurable w.r.t. $(t, a) \in D$ for all $u \in U$ and $y \in H$. For a.a. $(t, a) \in D$, ℓ is C^1 w.r.t. u and y , and there exists a modulus of continuity for ℓ , ℓ_y and ℓ_u , on bounded sets of U and the whole space H , independent of (t, a) .
- (H10) The functions $h_0, h_1 : H \rightarrow \mathbb{R}$ are C^1 .

By Lemma 4.2 and the assumptions (H9)-(H10), we can consider the *reduced cost* defined by

$$(93) \quad J_R(u, y_0, y_1, f) := J(u, y[u, y_0, y_1, f]),$$

and conclude that J_R is C^1 . In addition, for all $(v, z_0, z_1, h) \in \mathcal{U} \times L^2(0, A; H) \times L^2(0, T; H) \times L^2(D; V^*)$ we have

$$(94) \quad \begin{aligned} DJ_R(u, y_0, y_1, f)(v, z_0, z_1, h) &= \int_0^T \int_0^A \ell_y(u, y) z + \ell_u(u, y) v da dt \\ &\quad + \int_0^A h'_0(y(T, a)) z(T, a) da + \int_0^T h'_1(y(t, A)) z(t, A) dt, \end{aligned}$$

where z is the unique solution of the following system, whose existence is guaranteed by Section 3,

$$(95) \quad \begin{cases} z_t + z_a + \mathcal{A}(u)z + (D_u \mathcal{A}(u)v)y = h & \text{in } L^2(D; V^*), \\ z(0, \cdot) = z_0 & \text{in } L^2(0, A; H), \\ z(\cdot, 0) = z_1 & \text{in } L^2(0, T; H). \end{cases}$$

Now, assuming y_0, y_1 and f to be given, we consider the following reduced optimal control problem

$$(96) \quad \min_{u \in K_{\mathcal{U}}} J_R(u), \quad \text{s.t. } G_i(u) \in K_y, \quad i = 1, \dots, r,$$

where

$$(97) \quad G_i(u)(t) := \int_0^A (\xi_i(a), y[u](t, a))_H da - M_i, \quad t \in [0, T], \quad i = 1, \dots, r,$$

and $K_y := \{\varphi \in C([0, T]) : \varphi(t) \leq 0, \quad \forall t \in [0, T]\}$. In this case, we have

$$(98) \quad DG_i(u)v = \int_0^A (\xi_i(a), z[v](\cdot, a))_H da,$$

where $z[v]$ solves the system (95) with zero initial conditions and zero r.h.s.

We can define the *reduced Lagrangian* as

$$(99) \quad L_R(u, \beta, \mu) := \beta J_R(u) + \sum_{i=1}^r \langle \mu_i, G_i(u) \rangle = \beta J_R(u) + \sum_{i=1}^r \int_0^T G_i(u)(t) d\mu_i(t),$$

where $\beta \in \{0, 1\}$ and $\mu = (\mu_i)_{i=1}^r \in (BV_T(0, T))^r$.

Let $G(u) := (G_1(u), \dots, G_r(u)) \in K_y^r$. The constraints are said to be qualified at $u \in K_{\mathcal{U}}$ if there exists $\varepsilon > 0$ such that, B denoting the unit ball of $C([0, T])^r$, and e.g. $R_{K_{\mathcal{U}}}(u)$ denoting the radial cone to $K_{\mathcal{U}}$ at u :

$$(100) \quad \varepsilon B \subset DG(u)[R_{K_{\mathcal{U}}}(u)] - R_{K_y^r}(G(u)).$$

Since K_y^r is a convex cone with nonempty interior, it is known [5, Chapter 3] that this holds if and only if there exists $v \in K_{\mathcal{U}}$ such that

$$(101) \quad G(u) + DG(u)(v - u) \in \text{int } K_y^r.$$

In (102) below, N denote the normal cone, see [5, Chapter 2]. By [5, Theorem 3.18, p. 155], we obtain the following optimality condition.

Theorem 4.5. *Given y_0, y_1 and f , if $u \in K_{\mathcal{U}}$ is a local solution of the reduced problem, then there exists $\beta \in \{0, 1\}$ and $\mu = (\mu_i) \in (BV_T(0, T))^r$, with $\beta + \|\mu\| > 0$, such that*

$$(102) \quad 0 \in \beta DJ_R(u) + \sum_{i=1}^r DG_i(u)^* \mu_i + N_{K_{\mathcal{U}}}(u),$$

and

$$(103) \quad \mu_i \geq 0, \quad \langle \mu_i, G_i(u) \rangle_{C([0, T])} = 0, \quad i = 1, \dots, r.$$

In addition $\beta = 1$ if the qualification condition (100) is satisfied.

Definition 4.6. We define the *Lagrangian* L of the optimal control problem (63)-(67) as

$$(104) \quad \begin{aligned} L(u, y, \beta, p, \psi_0, \psi_1, \mu) = & \beta J(u, y) + \langle f - (y_t + y_a) - \mathcal{A}(u)y, p \rangle_{L^2(D; V)} \\ & + (\psi_0, y_0 - y(0, \cdot))_{L^2(0, A; H)} + (\psi_1, y_1 - y(\cdot, 0))_{L^2(0, T; H)} \\ & + \sum_{i=1}^r \langle \mu_i, \int_0^A (\xi_i(a), y(t, a))_H da - M_i \rangle_{C([0, T])}, \end{aligned}$$

where $\beta \in \{0, 1\}$, $p \in L^2(D; V)$, $\psi_0 \in L^2(0, A; H)$, $\psi_1 \in L^2(0, T; H)$ and $\mu \in (BV_T(0, T))^r$.

As usual, the *costate* equation is derived by computing $L_y z = 0$. Given $z \in W(D)$ the directional derivative of L with respect to y in the direction of z is

$$(105) \quad \begin{aligned} L_y z = & \beta \int_0^T \int_0^A \ell_y z \, da dt - \int_0^T \int_0^A \langle z_t + z_a + \mathcal{A}(u)z, p \rangle_V da dt \\ & + \beta \int_0^A h'_0(y(T, a)) z(T, a) da + \beta \int_0^T h'_1(y(t, A)) z(t, A) dt \\ & - (\psi_0, z(0, \cdot))_{L^2(0, A; H)} - (\psi_1, z(\cdot, 0))_{L^2(0, T; H)} \\ & + \sum_{i=1}^r \langle \mu_i, \int (\xi_i(a), z(t, a))_H da \rangle_{C([0, T])}. \end{aligned}$$

In order to analyze the regularity of the costate p , assume that there exists $(\beta, p, \psi_0, \psi_1, \mu) \in \{0, 1\} \times L^2(D; V) \times L^2(0, A; H) \times L^2(0, T; H) \times (BV_T)^r$ such that $L_y(u, y, \beta, p, \psi_0, \psi_1, \mu)z = 0$ is satisfied for all $z \in W(D)$. We define the alternative costate as

$$(106) \quad p^1(t, a) := p(t, a) + \sum_{i=1}^r \mu_i(t) \xi_i(a).$$

By (105) and our assumption that $L_y z = 0$ for all $z \in W(D)$, we obtain for all $z \in C_0^\infty(D; V)$,

$$(107) \quad \begin{aligned} 0 = & \beta \int_D \ell_y z \, da dt - \int_D \langle z_t + z_a + \mathcal{A}(u)z, p^1 \rangle_V da dt \\ & + \sum_{i=1}^r \int_D \langle z_t + z_a, \xi_i(a) \rangle_V \mu_i(t) da dt + \sum_{i=1}^r \int_D \langle \mathcal{A}^*(u) \xi_i(a), z \rangle_V \mu_i(t) da dt \\ & + \sum_{i=1}^r \int_0^T \int_0^A (\xi_i(a), z(t, a))_H da d\mu_i(t). \end{aligned}$$

By the integration by parts formula (86), we obtain

$$(108) \quad \begin{aligned} 0 &= \beta \int_D \ell_y z \, da dt - \int_D \langle z_t + z_a, p^1 \rangle_V da dt - \int_D \langle \mathcal{A}^*(u) p^1, z \rangle_V da dt \\ &\quad - \sum_{i=1}^r \int_D \langle \dot{\xi}_i(a), z \rangle_V \mu_i(t) da dt + \sum_{i=1}^r \int_D \langle \mathcal{A}^*(u) \xi_i(a), z \rangle_V \mu_i(t) da dt. \end{aligned}$$

So in a weak sense p^1 satisfies,

$$(109) \quad - (p_t^1 + p_a^1) + \mathcal{A}^*(u) p^1 - \beta \ell_y + \sum_{i=1}^r \mu_i(t) \dot{\xi}_i(a) - \sum_{i=1}^r \mu_i(t) \mathcal{A}^*(u) \xi_i(a) = 0, \quad \text{in } L^2(D; V^*).$$

By our assumptions, we deduce that $p^1 \in W(D)$. Then, in particular the traces $p^1(\cdot, a) \in L^2(0, T; H)$ for $a \in [0, A]$ and $p^1(t, \cdot) \in L^2(0, A; H)$ for $t \in [0, T]$ make sense. Since $L_y z = 0$, returning to (105) and replacing p by $p^1 - \sum_{i=1}^r \mu_i \xi_i$, by Lemma 3.3 we obtain the costate equation in transposed form,

$$(110) \quad \begin{aligned} 0 &= \int_0^T (z(t, 0), p^1(t, 0) - \sum_{i=1}^r \xi_i(0) \mu_i(t))_H dt \\ &\quad - \int_0^T (z(t, A), p^1(t, A) - \sum_{i=1}^r \xi_i(A) \mu_i(t))_H dt \\ &\quad + \int_0^A (z(0, a), p^1(0, a) - \sum_{i=1}^r \xi_i(a) \mu_i(0))_H da \\ &\quad - \int_0^A (z(T, a), p^1(T, a) - \sum_{i=1}^r \xi_i(a) \mu_i(T))_H da \\ &\quad + \beta \int_0^A h'_0(y(T, a)) z(T, a) da + \beta \int_0^T h'_1(y(t, A)) z(t, A) dt \\ &\quad - \int_0^A (\psi_0(a), z(0, a))_H da - \int_0^T (\psi_1(t), z(t, 0))_H dt. \end{aligned}$$

Then, since z is arbitrary, we deduce that p^1 satisfies the following age-structured system

$$(111) \quad \begin{cases} -(p_t^1 + p_a^1) + \mathcal{A}^*(u) p^1 - \beta \ell_y + \sum_{i=1}^r \mu_i(t) \dot{\xi}_i(a) - \sum_{i=1}^r \mu_i(t) \mathcal{A}^*(u) \xi_i(a) = 0, \\ \hspace{25em} \text{in } L^2(D; V^*), \\ p^1(T, a) - \sum_{i=1}^r \xi_i(a) \mu_i(T) = \beta h'_0(y(T, a)), & \text{a.e. } a \in [0, A], \\ p^1(t, A) - \sum_{i=1}^r \xi_i(A) \mu_i(t) = \beta h'_1(y(t, A)), & \text{a.e. } t \in [0, T], \end{cases}$$

and the following relations hold

$$(112) \quad \begin{cases} \psi_0(a) = p^1(0, a) - \sum_{i=1}^r \xi_i(a) \mu_i(0), & \text{a.e. } a \in [0, A], \\ \psi_1(t) = p^1(t, 0) - \sum_{i=1}^r \xi_i(0) \mu_i(t), & \text{a.e. } t \in [0, T]. \end{cases}$$

We can conclude that given $\beta \in \{0, 1\}$ and $\mu \in (BV_T)^r$, by Section 3 there exists a unique solution $p^1 \in W(D)$ of (111). Then, we obtain that p, ψ_0 and ψ_1 giving by

$$(113) \quad \begin{cases} p(t, a) &= p^1(t, a) - \sum_{i=1}^r \mu_i(t) \xi_i(a), \\ p(T, a) &= \beta h'_0(y(T, a)), \\ p(t, A) &= \beta h'_1(y(t, A)), \\ \psi_0(a) &= p(0, a), \\ \psi_1(t) &= p(t, 0). \end{cases}$$

satisfy the condition $L_y(u, y, \beta, p, \psi_0, \psi_1, \mu)z = 0$.

Definition 4.7. We define the *Hamiltonian* H as

$$(114) \quad \begin{aligned} H : U \times V \times \mathbb{R} \times V &\rightarrow \mathbb{R} \\ (u, y, \beta, p) &\mapsto H(u, y, \beta, p) := \beta \ell(u, y) + \langle \mathcal{A}(u) y, p \rangle. \end{aligned}$$

By the above definition and our assumptions, we can conclude that

$$(115) \quad H_u(t, a, u, y, \beta, p)v = \beta \ell_u(t, a, u, y)v + \langle (\mathcal{A}_u(t, a, u)v)y, p \rangle.$$

By (94) and (98), we have

$$\begin{aligned}
 D_u L_R(u, \beta, \mu)v &= \beta D J_R(u)v + \sum_{i=1}^r \langle \mu_i, D G_i(u)v \rangle_{C([0,T])} \\
 &= \beta \int_0^T \int_0^A \ell_y(u, y) z[v] + \ell_u(u, y) v \, da \, dt \\
 &\quad + \beta \int_0^A h'_0(y(T, a)) z[v](T, a) \, da + \beta \int_0^T h'_1(y(t, A)) z[v](t, A) \, dt \\
 &\quad + \sum_{i=1}^r \int_0^T \int_0^A (\xi_i(a), z[v](t, a))_H \, da \, d\mu_i(t),
 \end{aligned}
 \tag{116}$$

where $z[v]$ solves (95) with zero r.h.s. If p satisfies (113), we obtain

$$D_u L(u, y, \beta, p, \psi_0, \psi_1, \mu)v = \int_0^T \int_0^A H_u(t, a, u(t, a), y(t, a), \beta, p(t, a)) v(t, a) \, da \, dt.
 \tag{117}$$

Therefore, we obtain the following optimality condition:

Theorem 4.8. *Given y_0, y_1 and f , let u be a local solution of the optimal control problem. Then, there exists $\beta \in \{0, 1\}$, $p \in L^2(D; V)$, $\psi_0 \in L^2(0, A; H)$, $\psi_1 \in L^2(0, T; H)$ and $\mu \in (BV_T(0, T))^r$, with $\mu \geq 0$, $\beta + \|\mu\| > 0$, and*

$$\int_0^T \int_0^A (\xi_i(a), y(t, a))_H \, da \, d\mu_i(t) = 0, \quad i = 1, \dots, r,
 \tag{118}$$

such that p , ψ_0 and ψ_1 satisfy (113) and

$$\int_0^T \int_0^A H_u(t, a, u(t, a), y(t, a), \beta, p(t, a)) (\hat{u}(t, a) - u(t, a)) \, da \, dt \geq 0, \quad \forall \hat{u} \in K_{\mathcal{U}}.
 \tag{119}$$

4.3. Adding birth processes. As in the previous section, we also analyze optimality conditions for an optimal control problem in which the state has a birth process depending on a fecundity rate. We consider the state system studied in Subsection 3.3:

$$\begin{cases} (y_t + y_a)(t, a) + \mathcal{A}(t, a)y(t, a) = f(t, a) & \text{in } L^2(D; V^*) \\ y(0, a) = y_0(a) & \text{in } L^2(0, A; H) \\ y(t, 0) = \int_0^A c(t, a)y(t, a) \, da & \text{in } L^2(0, T; H). \end{cases}
 \tag{120}$$

Following the lines of the above subsection, given $(\beta, p, \psi_0, \psi_1, \mu) \in \{0, 1\} \times L^2(D, V) \times L^2(0, A; H) \times L^2(0, T; H) \times (BV_T)^r$ such that, $L_y(u, y, \beta, p, \psi_0, \psi_1, \mu)z = 0$ for all $z \in W(D)$, we can define the alternative costate p^1 as in (106). We deduce in this case that the alternative costate p^1 satisfies the following system:

$$\begin{cases} -(p_t^1 + p_a^1) + \mathcal{A}^*(u)p^1 - \beta \ell_y + \sum_{i=1}^r \mu_i(t) [\dot{\xi}_i(a) - \mathcal{A}^*(u)\xi_i(a)] = \psi_1(t)c(t, a) \\ \hspace{25em} \text{in } L^2(D; V^*) \\ p^1(T, a) - \sum_{i=1}^r \xi_i(a)\mu_i(T) = \beta h'_0(y(T, a)), & \text{a.e. } a \in [0, A] \\ p^1(t, A) - \sum_{i=1}^r \xi_i(A)\mu_i(t) = \beta h'_1(y(t, A)), & \text{a.e. } t \in [0, T] \\ \psi_1(t) = p^1(t, 0) - \sum_i \xi_i(0)\mu_i(t). \end{cases}
 \tag{121}$$

In order to prove that the above system has a unique solution $(\psi, w) \in L^2(0, T; H) \times W(D)$, it is enough to study the following system

$$\begin{cases} -(w_t + w_a)(t, a) + \mathcal{A}^*(t, a)w(t, a) = \psi(t)c(t, a) \\ w(T, a) = 0, \\ w(t, A) = 0, \\ \psi(t) = w(t, 0). \end{cases}
 \tag{122}$$

We define the map $\Phi : L^2(0, T; H) \rightarrow L^2(0, T; H)$, as $\Phi(\psi) := w(\cdot, 0)$, where $w \in W(D)$ is the solution of the following system, whose existence is guaranteed by Section 3,

$$(123) \quad \begin{cases} -(w_t + w_a)(t, a) + \mathcal{A}^*(t, a)w(t, a) = \psi(t)c(t, a) \\ w(T, a) = 0, \\ w(t, A) = 0. \end{cases}$$

Let $\gamma > 0$, we define the auxiliary map $\bar{w}(t, a) := e^{-\gamma t}w(T - t, A - a)$ which solves the following system

$$(124) \quad \begin{cases} (\bar{w}_t + \bar{w}_a)(t, a) + \gamma \bar{w} + \mathcal{A}^*(T - t, A - a)\bar{w}(t, a) = e^{-\gamma t}\psi(T - t)c(T - t, A - a) \\ \bar{w}(0, a) = 0, \\ \bar{w}(t, 0) = 0. \end{cases}$$

Then, applying the first equation to \bar{w} , by Lemma 3.3 we obtain

$$(125) \quad \begin{aligned} \frac{1}{2} \int_0^T (\|\bar{w}(t, A)\|_H^2 - \|\bar{w}(t, 0)\|_H^2) dt &= \frac{1}{2} \int_0^A (\|\bar{w}(0, a)\|_H^2 - \|\bar{w}(T, a)\|_H^2) da \\ &\quad - \gamma \int_D \|\bar{w}(t, a)\|_H^2 da dt - \int_D \langle \mathcal{A}(T - t, A - a)\bar{w}(t, a), \bar{w}(t, a) \rangle da dt \\ &\quad + \int_D (e^{-\gamma t} \Psi(T - t)c(T - t, A - a), \bar{w}(t, a))_H da dt. \end{aligned}$$

By our assumption (57), Cauchy-Schwarz and Young's inequalities we have

$$(126) \quad \begin{aligned} \frac{1}{2} \|\bar{w}(\cdot, A)\|_{L^2(0, T; H)}^2 &\leq (\lambda - \gamma) \|\bar{w}\|_{L^2(D; H)}^2 - \alpha \|\bar{w}\|_{L^2(D; V)}^2 \\ &\quad + \|e^{-\gamma t} \Psi(t)c(t, a)\|_{L^2(D; H)} \|\bar{w}\|_{L^2(D; H)} \\ &\leq \frac{\varepsilon}{2} \bar{C} \|\Psi\|_{L^2(0, T; H)}^2 + \left(\frac{1}{2\varepsilon} + \lambda - \gamma\right) \|\bar{w}\|_{L^2(D; H)}^2. \end{aligned}$$

Taking $\varepsilon > 0$ such that $\varepsilon \bar{C} < 1$ and $\gamma > 0$ such that $\frac{1}{\varepsilon} - 2\gamma + 2\lambda < 0$, by a fixed-point argument we can conclude that there exists a unique solution $(\psi, w) \in L^2(0, T; H) \times W(D)$ of (122).

Finally, we conclude that if we consider a birth condition given by (120), we obtain an optimality condition as in Theorem 4.8 where p^1 and ψ_1 satisfy system (122) instead of (111)-(112).

4.4. Fokker Planck type formulation. In order to give an example of application, in this section we explain how our framework includes the most usual formulation considered in the literature for dynamics of populations.

We consider a bounded domain $\Omega \subset \mathbb{R}^n$ with smooth enough boundary Γ . We set $V := H^1(\Omega)$ and $H := L^2(\Omega)$. Let $b : D \times \Omega \times U \rightarrow \mathbb{R}^n$ and $\hat{a} : D \times \Omega \rightarrow \mathbb{R}^{n \times n}$ be given maps. We consider the following model, that can be thought as a Fokker-Planck type equation for age-dependent dynamics.

$$(127) \quad \begin{cases} (y_t + y_a)(t, a, x) + \operatorname{div}(\Phi(t, a, x, u(t, a))) = f(t, a, x), & \text{in } D \times \Omega \\ \Phi(t, a, x, u(t, a)) \cdot n(x) = 0, & \text{in } D \times \Gamma, \\ y(0, a, x) = y_0(a, x), & \text{in } [0, A] \times \Omega \\ y(t, 0, x) = y_1(t, x), & \text{in } [0, T] \times \Omega \end{cases}$$

where

$$(128) \quad \Phi_i(t, a, x, u) := y(t, a, x)b_i(t, a, x, u) - \sum_{j=1}^n \frac{\partial(\hat{a}_{ij}(t, a, x)y(t, a, x))}{\partial x_j}, \quad i = 1, \dots, n.$$

The zero normal flux boundary condition can be stated as

$$(129) \quad y(t, a, x)b(t, a, x, u(t, a)) \cdot n(x) = \sum_{i,j=1}^n \frac{\partial(\hat{a}_{ij}(t, a, x)y(t, a, x))}{\partial x_j} n_i(x), \quad x \in \Gamma.$$

Then, the variational formulation for this equation can be written as

$$(130) \quad \int_D \int_\Omega (y_t + y_a)v - \Phi \cdot \nabla v \, dx da dt = \int_D \int_\Omega f v \, dx da dt, \quad \forall v \in L^2(D; V).$$

We make the following assumptions:

- (1) The functions \hat{a}_{ij} are measurable w.r.t. (t, a, x) and C^1 w.r.t. x , for all $i, j = 1, \dots, n$.
- (2) \hat{a}_{ij} and $\frac{\partial \hat{a}_{ij}}{\partial x_k}$ are essentially bounded for all $i, j, k = 1, \dots, n$, and for some $\alpha_0 > 0$

$$(131) \quad \sum_{i,j} \hat{a}_{ij}(t, a, x) \zeta_i \zeta_j \geq \alpha_0 |\zeta|^2, \quad \forall \zeta \in \mathbb{R}^n, \quad (t, a, x) \in D \times \Omega.$$

- (3) b is measurable w.r.t. (t, a, x) for all $u \in U$, and is C^1 w.r.t. $u \in U$ for a.a. $(t, a, x) \in D \times \Omega$.
- (4) For b and b_u there exists a modulus of continuity w.r.t. u , on bounded sets of U , which is independent of (t, a) .
- (5) b and b_u are bounded on bounded sets of U .

Under the above assumptions, we define $\mathcal{A}(t, a, x, u)$ by

$$(132) \quad \langle \mathcal{A}(t, a, u)y, v \rangle := \int_\Omega y(x)b(t, a, x, u) \cdot \nabla v(x) dx + \sum_{i,j=1}^n \int_\Omega \frac{\partial(\hat{a}_{ij}(t, a, x)y(x))}{\partial x_j} \frac{\partial v(x)}{\partial x_i} dx,$$

for $y, v \in V$. It is clear that \mathcal{A} satisfies (H5)-(H7).

In addition, we consider state constraints given by

$$(133) \quad \int_0^A \int_\Omega \xi_i(a, x) \cdot y(t, a, x) \, dx da \leq M_i, \quad i = 1, \dots, r, \quad \forall t \in [0, T],$$

where $\xi_i(a, x) := \xi_i(a)(x)$. In conclusion, all the results of this section apply for this formulation.

For the sake of simplicity, we derive the adjoint equation for the case of a linear cost function given by:

$$(134) \quad J(u, y) := \int_D \int_\Omega \ell(t, a, x)y(t, a, x) dx da dt + \int_D \int_\Gamma g(t, a, x)y(t, a, x) dx da dt.$$

Computing the derivative of the Lagrangian function with respect to the state variable, as in the above subsection, we obtain the following equation for the alternative costate p^1 :

$$(135) \quad \begin{cases} - (p_t^1 + p_a^1) + b(u) \cdot \nabla_x p^1 + \sum_{i,j=1}^n \hat{a}_{ij} \frac{\partial^2 p^1}{\partial x_i \partial x_j} - \beta \ell + \sum_{i=1}^r \mu_i(t) \dot{\xi}_i(a), \\ \quad - \sum_{k=1}^r \mu_k(t) \left[b(u) \cdot \nabla_x \xi_k + \sum_{i,j=1}^n \hat{a}_{ij} \frac{\partial^2 \xi_k}{\partial x_i \partial x_j} \right] = 0, & \text{in } D \times \Omega \\ \sum_{i,j=1}^n \hat{a}_{ij} \frac{\partial p^1}{\partial x_i} n_j(x) = g, & \text{in } D \times \Gamma \\ p^1(T, a, x) - \sum_{i=1}^r \xi_i(a) \mu_i(T) = 0, & \text{in } [0, A] \times \Omega \\ p^1(t, A, x) - \sum_{i=1}^r \xi_i(A, x) \mu_i(t) = 0, & \text{in } [0, T] \times \Omega. \end{cases}$$

In particular, if we take $b(t, a, x, u) = u$, $\hat{a}_{ij} = \delta_{ij}$ ($\delta_{ij} = 1$ if $i = j$ and 0 otherwise) and $\xi_i \equiv 1$, we obtain the following controlled system

$$(136) \quad \begin{cases} (y_t + y_a)(t, a, x) + \operatorname{div}_x(u(t, a)y(t, a, x)) - \Delta_x y(t, a, x) = f(t, a, x), & \text{in } D \times \Omega \\ u(t, a)y(t, a, x) \cdot n(x) = \nabla_x y(t, a, x) \cdot n(x), & \text{in } D \times \Gamma, \\ y(0, a, x) = y_0(a, x), & \text{in } [0, A] \times \Omega \\ y(t, 0, x) = y_1(t, x), & \text{in } [0, T] \times \Omega \\ \int_0^A \int_\Omega y(t, a, x) \, dx da \leq M & \forall t \in [0, T]. \end{cases}$$

In this case the state constraint represents a bound for the total population at any time. Note that our general framework allows us to consider systems like (136). This is one of the main differences between our work and most of the literature, where the control only

appears as a source term or by multiplying the state function, but not within a divergence term as in this example.

For this particular example, the equation of the alternative costate is given by:

$$(137) \quad \begin{cases} -(p_t^1 + p_a^1) + u \cdot \nabla_x p^1 + \Delta_x p^1 = \beta \ell, & \text{in } D \times \Omega \\ \nabla_x p^1(t, a, x) \cdot n(x) = g(t, a, x), & \text{in } D \times \Gamma \\ p^1(T, a) - \mu(T) = 0, & \text{in } [0, A] \times \Omega \\ p^1(t, A) - \mu(t) = 0, & \text{in } [0, T] \times \Omega. \end{cases}$$

5. EXISTENCE OF OPTIMAL CONTROLS

In the general case it is not easy to obtain results about the existence of optimal controls. Due to the bilinearity in the state equation, with respect to state and control, passing to the limit in this equation when we only have weak convergence results is not obvious. However, under particular assumptions, such as that the image spaces of u and y are \mathbb{R} , in [1] the authors achieved to prove the existence of minimizers. It is not difficult to extend this kind of results to our framework.

Nevertheless, in this section we present a modified control problem, for which it is possible to show the existence of solutions.

We consider $L^\infty(0, T; \mathbb{R})$ as the set of controls, i.e. the control functions only depend on t , transforming the control into a state. The map $\mathcal{A} : D \times \mathbb{R} \rightarrow L(V, V^*)$ is defined as

$$(138) \quad \mathcal{A}(t, a, u)y := \mathcal{A}_0(t, a)y + uy.$$

We assume $f \equiv 0$ and $\ell(u, y) := \ell_0(u) + \|y\|_H^2$, where ℓ_0 is uniformly continuous. We assume that (H5)-(H8) hold.

Now, we consider the following optimal control problem, where $v \in L^\infty(0, T; \mathbb{R})$ is the control and (u, y) the state,

$$(139) \quad \text{Min}_{v, y, u} \int_D [\ell_0(u) + \|y\|_H^2] da dt$$

s.t.

$$(140) \quad \begin{aligned} [p] \quad & y_t + y_a + \mathcal{A}(u)y = 0, & \text{in } D, \\ [\psi_0] \quad & y(0, a) = y_0(a), & a \in [0, A], \\ [\psi_1] \quad & y(t, 0) = y_1(t) & t \in [0, T], \\ [\mu_i] \quad & \int_0^A (\xi_i(a), y(t, a))_H da - M_i \leq 0, \quad i = 1, \dots, r, & \forall t \in [0, T], \\ [q] \quad & \dot{u}(t) = v(t), & t \in [0, T], \\ [\psi] \quad & u(0) = u_0 \\ [\eta] \quad & \frac{1}{2}|u(t)|^2 \leq M_u, & \forall t \in [0, T], \\ & |v(t)| \leq M_v, & \text{a.a. } t \in [0, T]. \end{aligned}$$

In this case we also write in the column on the left the associated multipliers.

5.1. Existence of solution.

Theorem 5.1. *Under the above assumptions, there exists an optimal control v in $L^\infty(0, T; \mathbb{R})$.*

Proof. Let $(v^n) \subset L^\infty(0, T; \mathbb{R})$ be a minimizing sequence. Since it is bounded by M_v , there exists a subsequence, still denoted (v^n) that converges to some \bar{v} in the $*$ -weak topology of $L^\infty(0, T; \mathbb{R})$. Let (u^n) and (y^n) be the associated states. The sequence (u^n) is uniformly bounded and equicontinuous on $[0, T]$, then by the Ascoli-Arzelá theorem, there exists a subsequence, still denoted (u^n) , that converges uniformly to $\bar{u} \in C([0, T])$. Finally, since (y^n) is bounded in $W(D)$, there exists a subsequence that weakly converges to some $\bar{y} \in W(D)$, and then $y^n \rightharpoonup \bar{y}$ in $L^2(D; V)$.

Since the characteristic functions $\chi_{[0,t]}$ belongs to $L^1(0, T)$ for all $t \in [0, T]$, and (v^n) $*$ -weakly converges to \bar{v} in $L^\infty(0, T; \mathbb{R})$, we have

$$(141) \quad u^n(t) = u_0 + \int_0^t v^n(t) dt \rightarrow u_0 + \int_0^t \bar{v}(t) dt, \quad \forall t \in [0, T].$$

We can conclude that \bar{u} is the state associated to \bar{v} . Now we want to prove that \bar{y} is the state associated to \bar{u} .

Let $z \in W(D)$, then by Lemma 3.3,

$$(142) \quad \begin{aligned} \Delta_1 &:= \int_D \langle y_t^n + y_a^n + \mathcal{A}_0 y^n + u^n y^n, z \rangle_V da dt \\ &= - \int_D \langle z_t + z_a, y^n \rangle_V da dt + \int_D \langle \mathcal{A}_0 y^n, z \rangle_V da dt \\ &\quad + \int_D u^n(t) (y^n(t, a), z(t, a))_H da dt \\ &\quad + \int_0^A [(y^n(T, a), z(T, a))_H - (y^n(0, a), z(0, a))_H] da \\ &\quad + \int_0^T [(y^n(t, A), z(t, A))_H - (y^n(t, 0), z(t, 0))_H] dt. \end{aligned}$$

We have

$$(143) \quad \begin{aligned} \int_D u^n (y^n, z)_H - \bar{u}(\bar{y}, z)_H da dt &\leq \int_D [(u^n - \bar{u})(y^n, z)_H da dt + \bar{u}(y^n - \bar{y}, z)_H] da dt \\ &\leq \sup_{t \in [0, T]} |u^n(t) - \bar{u}(t)| \int_D |(y^n, z)_H| da dt \\ &\quad + \int_D (y^n - \bar{y}, \bar{u} z)_H da dt. \end{aligned}$$

Since (u^n) converges uniformly to \bar{u} and (y^n) converges weakly in $L^2(D; H)$ to \bar{y} , we obtain that the r.h.s. of the above equation converges to zero.

On the other hand, since $y^n \rightharpoonup \bar{y}$ in $L^2(D; V)$, we have

$$(144) \quad \int_D \langle z_t + z_a, y^n - \bar{y} \rangle_V da dt \rightarrow 0$$

and

$$(145) \quad \int_D \langle \mathcal{A}_0 (y^n - \bar{y}), z \rangle_V da dt = \int_D \langle \mathcal{A}_0^* z, y^n - \bar{y} \rangle_V da dt \rightarrow 0.$$

Finally, by Lemma 3.3, the trace maps are continuous in the weak topology, then by (142)-(144) we can conclude that \bar{y} is the state associated to \bar{u} .

Now we have to prove that the state constraints are satisfied. For all $i = 1, \dots, r$ and $y \in W(D)$, we know that the map

$$(146) \quad t \in [0, T] \mapsto \int_0^A (\xi_i(a), y(t, a))_H da - M_i,$$

is continuous. Then, in order to obtain for all $t \in [0, T]$

$$(147) \quad \int_0^A (\xi_i(a), y(t, a))_H da - M_i \leq 0,$$

it is enough to prove the above inequality for a.a. $t \in [0, T]$. Assume there exists a measurable set $E \subset [0, T]$ with positive measure such that

$$(148) \quad \int_0^A (\xi_i(a), \bar{y}(t, a))_H da - M_i > \varepsilon, \quad \forall t \in E.$$

Let $\varphi := \chi_E$, it is clear that $\varphi \xi_i \in L^2(D; H)$, then

$$(149) \quad \int_D (\varphi(t) \xi_i(a), \bar{y})_H da dt - M_i > \varepsilon |E|.$$

Since $y^n \rightharpoonup \bar{y}$ in $L^2(D; H)$ we obtain

$$(150) \quad \int_D (\varphi(t) \xi_i(a), \bar{y} - y^n)_H da dt \rightarrow 0.$$

The state constraints are satisfied by y^n for all $n \in \mathbb{N}$, then we obtain a contradiction with (149). We can conclude that \bar{y} is the state associated with \bar{u} and satisfies the state constraints.

By the definition of the cost function, and the uniform convergence of (u^n) to \bar{u} and the weak convergence of (y^n) to \bar{y} , we can conclude that \bar{v} is an optimal control with associated states (\bar{u}, \bar{y}) . \square

5.2. Relationship with the original problem. The idea of this section is to show the relationship between the above optimal control problem and the original problem (63)-(68), described in Section 4, when we consider controls only depending on time. The intuition tells us that when the constraint for v is not active, the optimality condition for this problem is similar to the optimality condition obtained in Theorem 4.8.

We denote \tilde{L} the Lagrangian defined as:

$$(151) \quad \begin{aligned} \tilde{L}(v, u, y, \beta, p, \psi_0, \psi_1, \mu, q, \psi, \eta) &= L(u, y, \beta, p, \psi_0, \psi_1, \mu) + \langle q, v - \dot{u} \rangle \\ &\quad + \psi(u_0 - u(0)) + \int_0^T [\frac{1}{2}|u(t)|^2 - M_u] d\eta(t), \end{aligned}$$

where L is the Lagrangian of the original problem defined in (104). We define the Hamiltonian of this problem as

$$(152) \quad \tilde{H}(t, v, u, y, \beta, p, q) := \int_0^A H(t, a, u, y(a), \beta, p(a)) da + qv,$$

for $t \in [0, T]$, $\beta \in \{0, 1\}$, $v, u, q \in \mathbb{R}$ and $y, p \in L^2(0, A; H)$. It is clear that for $(u, y, \beta, p, \psi_0, \psi_1, \mu) \in \mathcal{U} \times L^2(D; V) \times \{0, 1\} \times L^2(D; V) \times L^2(0, A; H) \times L^2(0, T; H) \times (BV_T)^r$ we have

$$(153) \quad L_u(u, y, \beta, p, \psi_0, \psi_1, \mu) \hat{u} = \int_0^T \tilde{H}_u(t, v(t), u(t), y(t, \cdot), \beta, p(t, \cdot), q(t)) \hat{u}(t) dt.$$

Then, the costate q satisfies

$$(154) \quad q = \tilde{L}_v, \quad \text{and} \quad dq = -\tilde{H}_u dt - u d\eta.$$

Assume that $(\bar{v}, \bar{u}, \bar{y})$ is a solution of the problem (139)-(140). Then, there exists $(\beta, p, \psi_0, \psi_1, \mu, q, \psi, \eta)$ such that p is solution of the costate equation (113), q satisfies the above equation, $\mu \geq 0$, $\eta \geq 0$ and $\beta + \|\mu\| + \|\eta\| > 0$. In addition,

$$(155) \quad \text{supp}(d\eta) \subset \{t \in [0, T] : |u(t)|^2 = 2M_u\},$$

and

$$(156) \quad \int_0^T \int_0^A (\xi_i(a), y(t, a))_H da d\mu_i(t) = 0, \quad i = 1, \dots, r.$$

Finally, we have

$$(157) \quad q(v - \bar{v}) \geq 0, \quad \forall v \in L^\infty(0, T; \mathbb{R}) : |v(t)| \leq M_v.$$

Now, assume that there exists an interval $(t_1, t_2) \subset (0, T)$ where the constraint for \bar{v} is not active, i.e. $|\bar{v}(t)| < M_v$, for $t \in (t_1, t_2)$. Then, in this interval we obtain $q \equiv 0$, which implies $dq \equiv 0$ in (t_1, t_2) . We have for all $\hat{u} \in W^{1,\infty}(t_1, t_2; \mathbb{R})$,

$$(158) \quad 0 = \int_{t_1}^{t_2} \tilde{H}_u(t, v(t), u(t), y(t, \cdot), \beta, p(t, \cdot), q(t)) \hat{u}(t) dt + \int_{t_1}^{t_2} \bar{u}(t) \hat{u}(t) d\eta(t).$$

If we define the following set

$$(159) \quad \tilde{K}_{\mathcal{U}} := \{u \in W^{1,\infty}(0, T; \mathbb{R}) : u(t)^2 \leq 2M_u, t \in [0, T]\},$$

since (155) is satisfied, we can conclude that

$$(160) \quad \int_0^T \bar{u}(t)(u(t) - \bar{u}(t)) d\eta(t) \leq 0, \quad \forall u \in \tilde{K}_{\mathcal{U}}.$$

Combining the above equation and taking $\hat{u} := u - \bar{u}$ in (158), we conclude that

$$(161) \quad \int_{t_1}^{t_2} \tilde{H}_u(t, v(t), u(t), y(t, \cdot), \beta, p(t, \cdot), q(t))(u(t) - \bar{u}(t)) dt \geq 0, \quad \forall u \in \tilde{K}_{\mathcal{U}},$$

which is related to the optimality condition (119) obtained in Theorem 4.8, taking into account (117) and (153).

APPENDIX A. STANDARD PARABOLIC CASE

In this section we present classical results about the standard parabolic equation. See for instance [21, Chapter 3] and [18, Chapter 3].

Let $f \in L^2(0, T; V^*)$ and $z_0 \in H$, we consider the following system

$$(162) \quad \dot{z}(t) + \mathcal{A}(t)z(t) = f(t) \quad \text{in } L^2(0, T; V^*), \quad z(0) = z_0 \quad \text{in } H,$$

where $\mathcal{A} : [0, T] \rightarrow L(V, V^*)$ is a measurable function such that:

- (1) There exists $c > 0$ such that $\|\mathcal{A}(t)v\|_{V^*} \leq c\|v\|_V$, for all $v \in V$ and a.a. $t \in [0, T]$.
- (2) *Semicoercivity*: There exist $\alpha > 0$ and $\lambda \geq 0$ such that for all $v \in V$ and a.a. $t \in [0, T]$, $\langle \mathcal{A}(t)v, v \rangle_V \geq \alpha\|v\|_V^2 - \lambda\|v\|_H^2$.

Proposition A.1. *Under the above assumptions, there exists a unique solution $z \in W(0, T)$ of (162). In addition, there exists $C > 0$ such that*

$$(163) \quad \max_{t \in [0, T]} \|z(t)\|_H \leq Ce^{2\lambda T} [\|z_0\|_H + \|f\|_{L^2(0, T; V^*)}],$$

$$\|z\|_{W(0, T)} \leq C [\|z_0\|_H + \|f\|_{L^2(0, T; V^*)}].$$

Now, we add the following assumptions: $\mathcal{A}(t) = \mathcal{A}_0(t) + \mathcal{A}_1(t)$ where

- (3) $\mathcal{A}_0(t) : V \rightarrow V^*$ symmetric and continuously differentiable w.r.t. t , and there exists $\alpha_0 > 0$ such that $\langle \mathcal{A}_0(t)v, v \rangle \geq \alpha_0\|v\|_V^2$, for all $v \in V$ and a.a. $t \in [0, T]$.
- (4) $\mathcal{A}_1(t) : V \rightarrow V^*$ measurable with range in H and there exists $c_1 > 0$ such that $\|\mathcal{A}_1(t)v\|_H \leq c_1\|v\|_V$, for all $v \in V$ and a.a. $t \in [0, T]$.

The following theorem is known as the second parabolic estimates. We note that under different assumption, in [3] the author obtain a more general result.

Theorem A.2. *Assume $f \in L^2(0, T; H)$, $z_0 \in V$, and the assumptions (1)-(4) hold. Then, the solution z of the system (162) satisfies $z \in L^\infty(0, T; V)$, $\dot{z} \in L^2(0, T; H)$ and there exists $C > 0$ such that*

$$(164) \quad \|z\|_{L^\infty(0, T; V)} + \|\dot{z}\|_{L^2(0, T; H)} \leq C [\|z_0\|_V + \|f\|_{L^2(0, T; H)}].$$

Remark A.3. We can prove that $z(t) \in V$ for all $t \in [0, T]$ and $\sup_{t \in [0, T]} \|z(t)\|_V \leq C [\|z_0\|_V + \|f\|_{L^2(0, T; H)}]$. In fact, $z \in W(0, T) \subset C(0, T; H)$ and $z(t) \in V$ for a.a. $t \in [0, T]$. Consider $\varepsilon > 0$ and define $z^\varepsilon := \frac{1}{\varepsilon} \int_{T-\varepsilon}^T z(t) dt$.

By the previous proposition, since $z \in L^\infty(0, T; V)$ we obtain that $z^\varepsilon \in V$ and

$$(165) \quad \|z^\varepsilon\|_V \leq C [\|z_0\|_V + \|f\|_{L^2(0, T; H)}].$$

Since V is a Hilbert space, considering a sequence of $\varepsilon_n \downarrow 0$, we can take a subsequence, still denoted (z^{ε_n}) , that weakly converges to $\bar{z} \in V$. On the other hand, since $z \in C(0, T; H)$ we have that $\|z^{\varepsilon_n} - z(T)\|_H \rightarrow 0$. We then obtain that for all $w \in H$, $\langle w, z^{\varepsilon_n} \rangle_V \rightarrow \langle w, \bar{z} \rangle_V = \langle w, z(T) \rangle_H$, and also $\langle w, z^{\varepsilon_n} \rangle_H \rightarrow \langle w, z(T) \rangle_H$. We deduce $\bar{z} = z(T) \in V$. The same arguments prove that $z(t) \in V$ for all $t \in [0, T]$.

REFERENCES

- [1] B. Ainseba, S. Anita, and M. Langlais. Optimal control for a nonlinear age-structured population dynamics model. *Electronic Journal of Differential Equations*, 28:1–9, 2002. [1](#), [2](#), [19](#)
- [2] B. Ainseba and M. Langlais. On a population dynamics control problem with age dependence and spatial structure. *Journal of Mathematical Analysis and Applications*, 248(2):455–474, 2000. [1](#)
- [3] C. Bardos. A regularity theorem for parabolic equations. *Journal of Functional Analysis*, 7(2):311–322, 1971. [22](#)
- [4] J. F. Bonnans and P. Jaisson. Optimal control of a parabolic equation with time-dependent state constraints. *SIAM J. Control and Optimization*, 48(7):4550–4571, 2010. [12](#)
- [5] J. F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems*. Springer Series in Operations Research. Springer-Verlag, New York, 2000. [14](#)
- [6] M. Brokate. Pontryagin’s principle for control problems in age-dependent population dynamics. *Journal of Mathematical Biology*, 23(1):75–101, Dec 1985. [1](#)
- [7] G. Di Blasio. Non-linear age-dependent population diffusion. *Journal of Mathematical Biology*, 8(3):265–284, Oct 1979. [1](#)

- [8] J. Droniou. Intégration et Espaces de Sobolev à Valeurs Vectorielles. Working paper or preprint, Apr. 2001. [4](#), [5](#)
- [9] X. Dupuis. Optimal control of leukemic cell population dynamics. *Math. Model. Nat. Phenom.*, 9(1):4–26, 2014. [1](#)
- [10] S. Faggian. Hamilton Jacobi equations arising from boundary control problems with state constraints. *SIAM Journal on Control and Optimization*, 47(4):2157–2178, 2008. [1](#)
- [11] G. Feichtinger, G. Tragler, and V. M. Veliov. Optimality conditions for age-structured control systems. *Journal of Mathematical Analysis and Applications*, 288(1):47–68, 2003. [1](#)
- [12] M. G. Garroni and M. Langlais. Age-dependent population diffusion with external constraint. *Journal of Mathematical Biology*, 14(1):77–94, Mar 1982. [1](#), [3](#)
- [13] M. E. Gurtin. A system of equations for age-dependent population diffusion. *Journal of Theoretical Biology*, 40:389–392, 1973. [1](#)
- [14] M. E. Gurtin and R. C. MacCamy. Non-linear age-dependent population dynamics. *Archive for Rational Mechanics and Analysis*, 54(3):281–300, Sep 1974. [1](#)
- [15] M. Krastanov, N. Ribarska, and T. Tsachev. A Pontryagin maximum principle for infinite-dimensional problems. *SIAM Journal on Control and Optimization*, 49(5):2155–2182, 2011. [1](#)
- [16] M. Langlais. A nonlinear problem in age-dependent population diffusion. *SIAM Journal on Mathematical Analysis*, 16(3):510–529, 1985. [1](#), [3](#)
- [17] M. Langlais. Large time behavior in a nonlinear age-dependent population dynamics problem with spatial diffusion. *Journal of Mathematical Biology*, 26(3):319–346, Jun 1988. [1](#)
- [18] J.-L. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications. Vol. I*. Springer-Verlag, New York, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181. [3](#), [22](#)
- [19] N. P. Osmolovskii and V. M. Veliov. Optimal control of age-structured systems with mixed state-control constraints. *Journal of Mathematical Analysis and Applications*, 455(1):396–421, 2017. [1](#)
- [20] A. Ouedraogo and O. Traore. Optimal control for a nonlinear population dynamics problem. *Portugaliae Mathematica*, 62(2):217–229, 2005. [1](#)
- [21] F. Tröltzsch. *Optimal control of partial differential equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels. [22](#)
- [22] G. F. Webb. *Theory of nonlinear age-dependent population dynamics / G.F. Webb*. M. Dekker New York, 1985. [1](#)

INRIA-SACLAY AND CENTRE DE MATHÉMATIQUES APPLIQUÉES, ECOLE POLYTECHNIQUE, 91128 PALAISEAU, FRANCE (FREDERIC.BONNANS@INRIA.FR) .

CIFASIS-CONICET-UNR, OCAMPO Y ESMERALDA, S2000EZP, ROSARIO, ARGENTINA (GIANATTI@CIFASIS-CONICET.GOV.AR).